Track wear-and-tear cost by traffic class: Functional form, zero output levels and marginal cost pricing recovery on the French rail network

Marc Gaudry, Emile Quinet

PARIS SCHOOL OF ECONOMICS
ECOLE D'ECONOMIE DE PARIS

WORKING PAPER N° 2009 - 32

Track wear-and-tear cost by traffic class: Functional form, zero output levels and marginal cost pricing recovery on the French rail network

Marc Gaudry

Emile Quinet

JEL Codes: C20, C34, C87, D24, L92
Keywords: rail track wear-and-tear, cost function, CES, trans-Log, generalized Box-Cox, zero sample values, maintenance cost allocation by traffic class, marginal cost by traffic class, power axle weight damage laws, cross-sectional data, rail line sections, France, marginal cost pricing

PARIS-JOURDAN SCIENCES ECONOMIQUES
LABORATOIRE D'ECONOMIE APPLIQUÉE - INRA

48, BD JOURDAN – E.N.S. – 75014 PARIS
TÉL. : 33(0) 1 43 13 63 00 – FAX : 33 (0) 1 43 13 63 10
www.pse.ens.fr

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE – ÉCOLE DES HAUTES ÉTUDES EN SCIENCES SOCIALES
ÉCOLE NATIONALE DES PONTS ET CHAUSSÉES – ÉCOLE NORMALE SUPÉRIEURE

Track Wear and Tear Cost by Traffic Class: Functional Form, Zero-Output Levels and Marginal Cost Pricing Recovery on the French Rail Network

by

Marc Gaudry [1,3]
Émile Quinet [2]

1 Agora Jules Dupuit (AJD), Département de sciences économiques
Université de Montréal
Montréal

2 Paris-Jourdan Sciences Économiques (PSE)
École Nationale des Ponts et Chaussées (ENPC)
Paris

3 Institut National de Recherche sur les Transports et leur Sécurité (INRETS)
Université Paris-Est
Marne-la-Vallée

The authors thank *Société Nationale des Chemins de Fer Français* (SNCF) for providing the exceptional data base making this study possible and Bryan Breguet as well as Cong-Liem Tran of *Université de Montréal* for research assistance. The *Fonds Québécois de Recherche sur la Nature et les Technologies* (FQRNT) and the Natural Sciences and Engineering Council of Canada (NSERCC) supported the development of econometric tools used in this paper, an exploratory version of which (Publication AJD-66, Agora Jules Dupuit, *Université de Montréal*, and W. P. 04-02, *Centre d'Enseignement et de Recherche en Analyse Socio-économique, École Nationale des Ponts et Chaussées*) was presented at the First Conference on Railroad Industry Structure, Competition and Investment at the *Institut D'Économie Industrielle* (IDEI), Toulouse, November 7-8, 2003, and at the Eleventh World Conference on Transport Research, University of California at Berkeley, June 24-28, 2007. Some of the results included in this new version were produced within the CATRIN (Cost Allocation of TRansport INfrastructure cost) project funded by the Sixth Framework Programme of the European Community in 2008-2009.

Université de Montréal
Agora Jules Dupuit, Publication AJD-130

Paris-Jourdan Sciences Économiques
Working Paper N° 2009-32

Université Paris-Est
Institut National de Recherche sur les Transports et leur Sécurité
Département économie et sociologie des transports, Working Paper DEST N° 15

Version of 13th September 2009

ABSTRACT

We address the issue of the allocation of railway track maintenance (wear-and-tear) costs to traffic output classes and consider a very general function relating maintenance cost **C** to a set of technical production characteristics **K** used to produce traffic output vector **T**. We neglect other rail cost categories, such as traffic control and track renewal.

The data base pertains to over 1 500 sections of the French rail infrastructure in 1999, representing about 90% of the total network of 30 000 km of lines in regular service. In addition to the maintenance cost **C**, it provides by track section 15 technical characteristics (both state **S** and quality **Q**) and 4 train traffic outputs **T**. Input prices, assumed to be uniform in space, disappear from the analysis, as in other national cross-sectional cases.

With database subsets of approximately 1 000 observations, several functional forms are tested: Linear, Log-Log, Trans-Log and Generalized Box-Cox. All are embedded in an unrestricted extension (U-GBC) of Khaled's seminal restricted Generalized Box-Cox (R-GBC) functional specification. The U-GBC architecture, compared with its 4 principal nested variants, turns out to be by far the most appropriate, in particular when some observed zero Traffic sample values are included ─an issue rather neglected previously in the literature.

It appears that several technical characteristics, such as maximum allowed speed and number of switches, are highly significant maintenance cost factors, which gives a hint that derived marginal costs are short term; also, the relation between maintenance costs and traffic is non linear and differs significantly by train category. Implications of different specifications for marginal infrastructure cost charges by traffic type are outlined.

Key words: *rail track wear-and-tear, cost function, CES, Trans-Log, Generalized Box-Cox, zero sample values, maintenance cost allocation by traffic class, marginal cost by traffic class, power axle weight damage laws, cross-sectional data, rail line sections, France, marginal cost pricing.*

SOMMAIRE

Ce texte traite de l'allocation des coûts d'entretien de l'infrastructure ferroviaire, entendus au sens de « coûts à périodicité annuelle », aux divers types de trafic. Sont exclus de l'analyse les coûts de circulation (contrôle du trafic) et de renouvellement des voies dits « coût de régénération ». La fonction de coût utilisée, très générale, relie le coût **C** à un ensemble de caractéristiques techniques **K** et à un vecteur de trafics produits **T.**

Les données utilisées proviennent d'une base de données relative à la France sur une décomposition en 1 500 segments environ d'un réseau ferroviaire de 30 000 km de lignes en service en 1999. On dispose par segment, en plus du coût d'entretien **C**, de leurs 15 principales caractéristiques techniques (tant d'état **S** que de qualité **Q**) et du trafic **T** décomposé en 4 catégories de service. Les prix, supposés constants à travers le pays, disparaissent de l'analyse, à l'instar d'autres exemples construits à partir de coupes transversales nationales.

À partir de sous-ensembles d'approximativement 1 000 observations, différentes spécifications de relations fonctionnelles sont testées et évaluées: Linéaire, Log-Log, Trans-Log et Box-Cox Généralisée. Elles sont toutes comprises dans un prolongement non contraint (U-GBC) de la formulation fondatrice Box-Cox Généralisée contrainte (R-GBC) de Khaled. La structure générale U-GBC, comparée aux 4 principales formes qui y sont emboîtées, ressort clairement comme préférable, en particulier lorsqu'on retient dans l'échantillon les observations nulles sur certains trafics ─un sujet plutôt négligé à ce jour dans la littérature.

Il appert que le coût dépend des caractéristiques techniques telles que la vitesse maximale autorisée sur le segment de ligne et le nombre de croisements de voies, ce qui laisse entendre, comme les caractéristiques techniques couvrent les plus importantes caractéristiques de l'équipement, que les coûts marginaux déduits des ajustements représentent correctement des coûts de court terme ; en outre, la relation entre le trafic et les coûts d'entretien est non-linéaire et varie par type de train. On étudie enfin les implications des diverses spécifications pour la tarification des infrastructures au coût marginal par type de trafic.

Mots clés: *voie ferrée, entretien, fonction de coût, CES, Trans-Log, Box-Cox Généralisée, observations nulles, classe de trafic, coût marginal d'entretien par catégorie de trafic, fonction de dommage des charges-essieux, données en coupe instantanée, France, sections de ligne ferroviaire, tarification au coût marginal.*

Journal of Economic Literature: C20, C34, C87, D24, L92

TABLE of CONTENTS

INDEX of TABLES

INDEX of FIGURES

1. The problem: form and the assignment of joint costs to multiple train classes[1]

In this paper, we address the problem of the allocation of railway track maintenance costs, which exclude both traffic control (operation) and regeneration (reconstruction) costs, to different traffic classes. Of long concern to railways and their regulators, such as Nördling (1886) who states:

> "*How should total expenses be split between passenger and freight services? A few engineers have admitted, without proof, that a passenger train caused the same expense per kilometre as a freight train. In Austria, as the passenger trains are 4 to 5 times lighter than the freight trains, this formula would imply that the gross ton from a passenger train is 4 to 5 times more expensive than the gross ton from a freight train. The most used formula in France treats a passenger-kilometre as equivalent to a ton-kilometre of freight.*" (Our translation from the French original),

the question remains difficult to this day because the relationships are expected to be highly non linear, as theory (Borts, 1960) or practice (Fuss *et al*., 1978) suggest, and as pointed out in classic textbooks (*e.g*. Wilson, 1980), regulatory documents (*e.g*. Hariton, 1984) and best practice surveys (*e.g*. Thomas, 2002). Oum & Waters (1996, 1997) give a good account of the rise and persistence of simple regressions, linear in both parameters and variables (LIN), still found to-day despite the general expectation of non linearity: due to the lack of detailed analytic accounting data on maintenance costs and to the fact that traffic data by user class are confidential[2] or just unavailable, analysts may be obliged to relate costs (often for each expense category), within a linear-in-parameters model, to a linearly specified "service unit" such as *ton-km*, *loaded car-km*, or *tons* (Martland, 2001), making *ad hoc* adjustments for differences in passenger and freight ton-km mix.

Within linear-in-parameter models of interest in this study, Log-Log (LL) forms are common, if only as special cases of the Trans-Log[3] (TL) form. The latter, long dominant in the analysis of total rail cost, as surveys make clear (*e.g.* Jara-Diaz, 1982; Tovar *et al*., 2003), has recently emerged (Bereskin, 2000) and spread (Link *et al*., 2008) in maintenance cost studies as the dominant form there as well. We analyze maintenance costs here by comparing these known forms, all nested within the unrestricted Generalized Box-Cox (U-GBC) version of Khaled's (1978) restricted (R-GBC) applications. The U-GBC was first applied to rail costs in the exploratory version of this paper (Gaudry & Quinet, 2003) and the results, to be summarized below, prompted Wheat (Wheat *et al*., 2009) to promote and coordinate[4] in 2008-2009 a series of Box-Cox (BC) applications with national data (Andersson, 2009a, 2009b; Gaudry & Quinet, 2009; Link, 2009; Marti *et al*., 2009) to which we will refer again as they contain comparisons of LL and BC forms.

In the next section, we outline the characteristics of our database. We then turn more formally to a methodological strategy on form, in a short section. Results and their implications for cost recovery under hypothetical marginal cost charging *régimes* are finally discussed.

---

[1] The authors thank *Société Nationale des Chemins de Fer Français* (SNCF) for providing the exceptional data base making this study possible and Bryan Breguet as well as Cong-Liem Tran of *Université de Montréal* for research assistance. The *Fonds Québécois de Recherche sur la Nature et les Technologies* (FQRNT) and the Natural Sciences and Engineering Council of Canada (NSERCC) supported the development of econometric tools used in this paper, an exploratory version of which (Gaudry & Quinet, 2003) was presented at the First Conference on Railroad Industry Structure, Competition and Investment at the *Institut D'Économie Industrielle* (IDEI), Toulouse, November 7-8, 2003. Some of the results included in this extensive new version were produced within the CATRIN (Cost Allocation of TRansport INfrastructure cost) project funded by the Sixth Framework Programme of the European Community.

[2] In the absence of explicit mathematical forms and parameters, the reader often has to infer the presence of non linearity from the results, for instance in Robert *at al*. (1997).

[3] This abbreviation of "transcendental in the logarithms of its arguments" was used in the first presentation of the function (Christensen *et al*., 1971), and the simplified Translog denomination in the second (Christensen *et al*., 1973).

[4] All carried out within the CATRIN (Cost Allocation of TRansport INfrastructure cost) project funded by the Sixth Framework Programme of the European Community.

2. The data: the French rail network sections in 1999

We benefit from a high quality data base pertaining to about 1 500 sections of the French rail infrastructure for which wear and tear costs **C**, traffic levels and composition by user class **T**, and other technical segment-specific information on the state **S** and quality **Q** of the segment, are available for 1999 (Quinet, 2002). These data, covering about 90% of the total French network of 30 000 km and described in Table 1, are of a quality sufficient to explore the issue of form and allow the data themselves, as opposed to the analysts' presumptions, to determine the best fitting one for the explanation of the maintenance cost.

Table 1. Nature of information available by rail line section for the French network

C	maintenance **Cost**, drawn from the analytical accounts of *Société Nationale des Chemins de Fer Français* (SNCF), encompasses maintenance costs allocated to track segments and represents 81 % of total maintenance expenditure: the rest consists of triages, intermodal freight platforms and service tracks, none of which can clearly be assigned to a given track segment. It also covers catenaries, signalling, tracks, rails, sleepers, ballast, culverts and "works of art" (*ouvrages d'art*, *i.e.* bridges and tunnels). It excludes traffic control and renewal (regeneration/reconstruction) costs except for some Large Maintenance Operations (*Opérations de Grand Entretien*, OGE) which are conventionally assigned to maintenance accounts but are close to renewals[5].
S	technical **State** variables such as the number of tracks (from 1 to 18), the number of switches, the type of control devices (automatic or not), the type of power (electrified or not), the length of the section.
Q	technical **Quality** variables such as the age of rails, the age of sleepers, the share of concrete (*vs* wood) sleepers and the maximum allowed speed in normal operation (in the absence of incidents or repairs in progress). Maximum allowed speed is effectively both a state and a quality technical factor.
T	train **Traffic**, measured by the number and average weight (tons) of 5 types of trains, namely long distance passenger (GL = *Grandes Lignes*), Île-de-France passenger (IdF), other[6] regional passenger (TER = *Trains Express Régionaux*), freight (F) and track servicing (HLP = *Haut-le-pied*) trains.

Due to various conditions imposed on variables of interest for our tests, we define 3 databases: (i) for the ***exploratory*** work, a LARGE database of 1146 observations extracted from the source; (ii) for the ***reference*** tests a MEDIUM database of 967 observations extracted from the large one; (iii) for ***special tests*** on the role of zero-output values, a SMALL database of 928 observations retained from the medium one. These three nested bases are of course very similar, as a comparison of Table 2 and Table 3 for total and average segment length, or of relative traffic indicators, shows: for instance, segments add up to 20 397 km in the large extract and to 18 402 km in the small one.

The important difference between the LARGE and MEDIUM databases is that the former includes some 18 high speed rail-only[7] links excluded from the latter. But the main difference between the three databases occurs between the MEDIUM and the SMALL ones due to the exclusion of observations containing some 0 traffic: considering only two kinds of traffic, total passenger (T1+T2+T3) and freight (T4), the exclusion of 39 segments where total passenger traffic equals zero reduces the sample to 928 observations.

[5] Renewal cost, incurred only every 20 or 30 years, should not be related to the current traffic but to a "proper" cumulative amount of traffic, measured in "equivalent-tons" (depending on the specific effects, if any, of the traffic classes), since the last renewal. Whence, yearly section repairs do not provide valuable information on the cost function for renewal expenses which raise issues of their own.

[6] The distinction between TER and IdF traffic deserves some explanation: IdF traffic is the local traffic around the extended Paris area (12 million inhabitants) and is mainly suburban while TER traffic corresponds to local traffic in other parts (50 million inhabitants) of France and is a mix of suburban (around large agglomerations) and rural traffic.

[7] High speed lines, representing about 1 500 km of total lines in 1999, are only used by high speed rail trains (TGV), but such trains also run on all classic electrified lines, albeit at lower speeds.

Table 2. Range of traffic and segment length values, LARGE database of 1146 observations

	Trains types						*Segments*
	GL	**TER**	**IdF**	**Freight**	**Other**	**Total**	***Exact length***
Average weight (tons)	534[1]	212	390	1 055	98	2 290	***17 798 m***
Number of trains	1 971	1 463	1 745	1 828	438	7 445	
Maximum	29 524	19 499	61 835	16 383	7 665	92 077	***157 924 .***
Minimum	0	92	0	0	0	372	***238 m***

[1] Includes average weight of TGV trains (615).

Table 3. Principal characteristics of the SMALL database of 928 observations

Variable	Average	Maximum	Minimum
Cost of maintenance			
cost per km (1999 francs)	464 801	7 604 163	480
State			
switches per segment	22	345	1
length of line segments (metres)	19 197	157 924	238
power type (electrified or not)	0,68	1,00	0,00
type of traffic control (automatic or not)	0,77	1,00	0,00
Quality			
age of rail (years)	26	92	4
age of sleepers (years)	27	92	4
maximal allowed speed (km/h)	127	220	60
share of concrete (*vs* wood) sleepers	0,58	1,00	0,00
Traffic indices by traffic category [(number of trains)x(average weight in gross tons)]			
T1 : long distance passenger trains (tons)	1 116 095	16 809 701	0
T2 : regional passenger trains (tons)	426 421	4 476 021	0
T3 : Ile-de-France passenger trains (tons)	878 439	32 778 126	0
T4 : freight trains (tons)	2 478 063	16 442 960	281

If it were desired to exclude all observations containing zero values on the basis of the 4 traffic service categories, the sample would be reduced further to only 208 observations. Hence, if traffic is required to be always positive, a 4-traffic model is not comfortably feasible: it is only possible with the first two database variants where observed null values of traffic variables are all retained, an issue to be addressed below.

Everything considered, our source SNCF database is relatively rich in technical (**S**, **Q**) and traffic (**T**) information, as compared to what is publicly available in some other country analyses, *e.g.* for Sweden and Finland (Johansson & Nilsson, 2004), or to what is reported in European surveys (*e.g.* Abrantes *et al.*, Table 5, 2007). It contains no information on input prices but we shall in any case assume that, in 1999, such prices were uniform throughout France, a reasonable assumption due to the fact that workers' wages are uniform across the SNCF, that electricity is supplied through a nation-wide contract and that other input prices are not likely to differ much among the 23 SNCF administrative regions due to strong horizontal coordination organized by maintenance expense category. In these conditions, the absence of input prices from all of our models only implicitly rescales their regression intercepts, as also occurs since Johansson & Nilsson (2004) in other single-year national cross-sectional analyses that make the same input price hypothesis for similar reasons (*e.g.* Andersson, 2009a, 2009b).

Our baseline view is that all regression models should in any case always have intercepts because forcing the adjustment to go through the origin can change regression signs. The use of Box-Cox transformations only strengthens this position by requiring intercepts if the form parameters are to be invariant to units of measurement of variables, as demonstrated by Schlesselman (1971).

3. An econometric methodology allowing for a censored dependent variable

To the extent that we are interested in rail costs, it is clear that innovations in cost functions were first applied to the analysis of total rail costs, typically with time series, and spread much later to the analysis of rail cost categories, principally with cross sections.

To explain rail cost C_y, regression models use as explanatory variables factors belonging to three principal groups: input prices $P \equiv (P_1, \ldots, P_n)$, traffic outputs $T \equiv (T_1, \ldots, \ldots T_I)$ and service characteristics or qualities $Q \equiv (Q_1, \ldots, Q_U)$; the regression may take one or more possible forms, all nested in the *Unrestricted Generalized Box-Cox* (U-GBC), to which we presently turn.

It will then be convenient afterwards to formulate the modeling framework quite generally and to select proper statistical tools to evaluate the performance of the principal competing forms, popular or not, applied in rail maintenance cost analysis.

3.1. The Generalized Box-Cox (GBC) idea and time-series explanations of total cost

The U-GBC is obtained by simply requiring that, in the *generalized flexible quadratic* form studied by Blackorby *et al.* (1977), the unspecified $f_i(X_i)$ and $f_j(X_j)$ functions be Box-Cox transformations (BCT), and by transforming the dependent variable in similar fashion:

$$C_y^{(\lambda_y)} = \beta_0 + \sum_k^r \beta_k \, X_k^{(\lambda_k)} + \sum_i^r \sum_j^r \beta_{ij} \, X_i^{(\lambda_i)} \, X_j^{(\lambda_j)} \qquad \text{(U-GBC)}$$

with the transformation of the strictly positive dependent or independent variables Var_v defined as:

$$Var_v^{(\lambda)} \equiv \begin{cases} \dfrac{(Var_v)^{\lambda} - 1}{\lambda} & , \quad if\ \lambda \neq 0, \\ \ln(Var_v) & , \quad if\ \lambda \to 0. \end{cases} \qquad \text{(BCT)}$$

The U-GBC form generalizes the R-GBC specifications applied by Khaled (1978) and Berndt & Khaled (1979) to the first right hand side term of their cost function $C = \{[h(p_1, \ldots, p_p)] \cdot [g(p_1, \ldots, p_p; T)]\}$, where T is a single-output vector and the vector of prices P appears twice. Their strategy consists in making the price term $h(p_1, \ldots, p_p)$ flexible with BCT while keeping the second term $g(p_1, \ldots, p_p; T)$ fixed. To give make explicit, and give due credit to, this seminal use of BCT in cost functions, we have to momentarily neglect, for expository reasons, the second term of the product and write the remaining simplified function as:

$$y^{(\lambda)} = \beta_0 + \sum_k^r \beta_k \, X_k^{(\lambda)} + \sum_i^r \sum_j^r \beta_{ij} \, X_i^{(\lambda)} \, X_j^{(\lambda)}, \qquad \text{(R-GBC)}$$

where the restrictions imposed on (U-GBC) are obvious. It suffices for our purposes that, firstly, one can obtain the Trans-Log (TL) case[8] from (R-GBC) when $(\lambda_C = \lambda_k = \lambda_{p_i} = \lambda_{p_j}) \to 0$:

[8] After imposing some restrictions on the coefficients and after absorbing in those pertaining to interaction terms the ½ coefficients of the formal TL. The TL is also obtained with the multiplication of functions effected in Berndt & Khaled.

$$\ln(y) = \beta_0 + \sum_k^r \beta_k \ln(X_k) + \sum_i^r \sum_j^r \beta_{ij} \ln(X_i) \ln(X_j) . \qquad \text{(TL)}$$

But, secondly, we note that, in addition to a TL case, Berndt and Khaled in fact estimated another restricted variant called the Generalized Box-Cox (GBC), obtainable by imposing the restrictions $[\hat{\lambda}_C ; (\hat{\lambda}_k = \hat{\lambda}_{p_i} = \hat{\lambda}_{p_j} = \hat{\lambda}_C/2)]$ on a price expression $h(p_1,...,p_p)$ of U-GBC form[9]. They compared this GBC formulation to the Generalized Square Root Quadratic (GSRQ) special case, specified as $[\lambda_C = 2 ; (\lambda_k = \lambda_{p_i} = \lambda_{p_j} = \lambda_C/2 = 1)]$, and to the Generalized Leontief (GL) other special case, specified as $[\lambda_C = 1 ; (\lambda_k = \lambda_{p_i} = \lambda_{p_j} = \lambda_C/2 = 1/2)]$, both introduced by Diewert (1971, 1973, 1974).

This approach, consisting in multiplying the partially flexible form component $h(p_1,...,p_p)$ by the fixed form component $g(p_1,..., p_p ; T)$, was applied to U.S. manufacturing output with annual data from 1947 to 1971. The results, requiring at most estimation of a single BCT, made it possible to reject the TL and the GSRQ but not the GL, when compared to the GBC.

On practical grounds then, and despite the fact that part of the cost function specification remained of untested fixed form, they demonstrated that the BCT makes it possible to *choose among many known competing second order fixed form approximations to an arbitrary cost function.* Their approach should have established the baseline of subsequent cost analyses, notably in transport where the TL form, applied straightforwardly to all variables at hand, became extremely popular.

A rail extension: multiple outputs and qualities. But the only proper transport following we could find is Borger (1992) who, in a study of yearly railroad operations in Belgium from 1950 to 1986, astutely extended the above Berndt-Khaled analysis by enriching the second term of the cost function, rewritten as $g(p_1,..., p_p ; T ; Q)$, where T now contained two outputs (passenger and freight), and each output was related to various network qualities Q called "operation characteristics". Variables found in **T** and **Q** vectors were regrouped in strictly positive "output aggregator" terms of fixed form with coefficients estimated independently.

Borger did not modify the basic multiplication of functions at the heart of the Berndt & Khaled cost formulation, but may have considered the prospect when he included the exogenous estimates: "Note that we did not estimate the parameters of the aggregator functions simultaneously with the GBC model. Although this would have been preferable from a theoretical perspective, the complexity of the GBC model forced us to use a simpler alternative. We therefore used the [log linear] aggregates [previously] constructed in De Borger (1991) as independent variables in the estimation" of GBC, TL and GL forms.

The Box-Tidwell patch of Trans-Log models with zero-output levels. Another special case of an application directly nested in the U-GBC form is the "Generalized Translog Multiproduct Cost Function" (GTMCF) proposed and discussed by Burgess (1974), Brown *et al.* (1979) and Caves *et al.* (1980b), and which consists in working directly on the TL specification. Caves *et al.* (1980a; 1985) explain railway (total) cost with this slightly modified TL form departing from the similar treatment of all explanatory variables (input prices and two outputs in this case). This modification is described in the former study: "This cost function has the same form as the Trans-Log except for output levels, where the Box-Cox metric is substituted for the natural log metric. This generalization permits the inclusion of firms with zero-output levels for some products": its

[9] Their particular GBC case is obtained by effectively solving for C_y after applying the U-GBC format to price term $h(p_1,...,p_p)$ and then multiplying the result by the $g(p_1,..., p_p ; T)$ term always assumed to be of fixed form.

implementation with freight and passenger traffic was required "since passenger service is zero for a substantial number of observations" (Caves *et al.,* 1985).

The use of the GTMCF did not spread much, as it "highly complicates the interpretation of parameters" (Tovar *et al.*, 2003, Footnote 9) and consequently makes the imperative calculation of elasticities still more burdensome, a point we address shortly. It amounts to a very marginal change of the TL because a single BCT is used on two train service variables: it is therefore the most limited Box-Tidwell[10] patch possible[11]. Why such parsimony? The note by Caves *et al.*, (1980a, Footnote 3) states that a generalization to more than one BCT "needlessly complicates estimation". As we shall see, transforming many traffic variables in a rail cost problem would often be a better choice, notwithstanding the real additional estimation complications, to which we turn shortly.

3.2. Multi-product cross-sectional models of rail maintenance cost

Studies of rail maintenance cost. The above mentioned studies of total rail cost making some use of BCT are all estimated from firm-wide time series data. But our specific interest is in studies of rail cost categories, such as maintenance cost, whether time series or pooled cross-sectional and time series. Studies of rail cost categories are relatively recent and predominantly use TL forms, for instance Bereskin (2000) or Sánchez (2000) with firm-level pooled data of an "aggregate" nature, and Link *et al.* (2008) for (a summary of) work done with detailed track-level data.

Although time series make it possible to maintain input price variables in the specifications, we could find none that also included indices of track quality. By contrast, almost all single-firm studies based on detailed rail track segment data are cross-sectional or estimated as cross-sectional data sets because the number of available years of data on track segments is extremely short, *e.g.* Johansson & Nilsson (2004). In these models, price variables normally disappear and one is left with multiple outputs in **T** and track qualities in **Q**.

Maintenance cost and the U-GBC form. The corresponding general U-GBC form is therefore:

$$C^{(\lambda_C)}=\alpha_0+\sum_{i=1}^{U}\alpha_i Q_i^{(\lambda_q)}+\sum_{j=1}^{I}\beta_j T_j^{(\lambda_p)}+\sum_{i=1}^{U}\sum_{j=1}^{U}\gamma_{ij} Q_i^{(\lambda_q)} Q_j^{(\lambda_q)}+\sum_{i=1}^{I}\sum_{j=1}^{I}\delta_{ij} T_i^{(\lambda_p)} T_j^{(\lambda_p)}+\sum_{i=1}^{U}\sum_{j=1}^{I}\zeta_{ij} Q_i^{(\lambda_q)} T_j^{(\lambda_p)}+u\ , \quad (0)$$

where, for simplicity, the regression coefficients have absorbed the ½ factor applied to all interaction sequences of the original Trans-Log and where, in addition to the required intercept, two vectors of explanatory variables $Q \equiv (Q_1, \ldots, Q_q, \ldots, Q_U)$ and $T \equiv (T_1, \ldots, T_p, \ldots T_I)$ are included and the parentheses again refer to Box-Cox transformations (BCT) of the strictly positive dependent or independent variables.

This assumption of strictly positive values implies that no sample value of a variable is ever equal to zero. In these conditions, the U-GBC naturally includes the nested Linear (LIN) and TL forms; and the TL is easily compared, by dropping all interaction terms, with the well known competing Constant Elasticity of Substitution (CES) or Cobb-Douglas Log-Log (LL) form previously in common use for railway production functions (*e.g.* Baumgartner, 1978).

[10] In the literature, applying one or more BCT only to *explanatory* variables is simply called a Box-Tidwell (Box & Tidwell, 1962) model because the *dependent* variable is not transformed as in proper Box-Cox (1964) models.

[11] We will address below the problem posed by this apparently innocuous practice, confirmed in an August 2003 correspondence with one of Caves' co-authors. The fact that, when $\lambda \neq 0$, the BCT is well defined for zero-output levels [the result equals $-1/\lambda$] does not mean that any vector of observations on a variable X_z with some zeroes may be transformed as if it were strictly positive: in the presence of some zero values, the BCT will not be invariant to the scale chosen for X_z. We discuss below remedies to this lacking shift [of $-1/\lambda$ when $X_{zt} = 0$] that re-establish "invariance to units of measurement", as it is commonly called.

In conformity with Table 3, consider that, in (0), the dependent variable ***C*** designates *the rail Cost per kilometre* to be explained, the $\boldsymbol{Q_q}$ refer to *track Qualities* (neglecting the State variables, almost all of which are binary) and the $\boldsymbol{T_p}$ to jointly produced *Traffic service* classes. Also neglect, without loss of generality, potentially interesting *a priori* restrictions that could be imposed on the α and β coefficients, notably in the Trans-Log special case[12], but retain the parsimonious and reasonable symmetry conditions on the matrices of γ, δ and ζ coefficients[13]. Finally, to simplify presentational notation without actual loss of generality, momentarily consider only as many λ_q as there are Qualities (q = 1, … , U) and only as many λ_p as there are Traffic service products (p = 1, … , I); and neglect specifications arising from interaction-specific combinations[14] of λ_q and λ_p BCT powers despite their expected relevance in practice and use below.

Our U-GBC form therefore allows multiple outputs in **T** and effectively replaces input prices of time series by qualities **Q**. It is considered here pragmatically, to quote Christensen *at al*. (1971) on the TL form, as a potentially valid "second order approximation to an arbitrary functional form" to be compared empirically to other popular forms in the absence of zero-output values.

Below, we estimate many such nested variants of U-GBC directly in their natural "level" format, including the TL, simply assuming that the SNCF management minimizes short-run maintenance costs for a track system of given technical characteristics ***K* ≡ *(S, Q)*.** We do not discuss whether the TL, either in cost *level* or in cost *share* format[15], is as adequate for monopolies[16] as it may be for competitive firms and do not use in the estimation input factor share equations which are neither necessary nor available.

3.3. Overcoming barriers to optimal form GBC estimation

Forgetting the baseline. We asked above why the powerful demonstration by Berndt & Khaled was so little applied elsewhere, and notably in transport. To start with, as implicit in the remarks by Caves *et al*. (1980a) and Borger (1992) quoted above, generality requires an algorithm to estimate the Box-Cox power parameters, respecting at the same time any desired constraints imposed on the regression coefficients. Despite its limitations, for instance on the possibility of testing sum constraints on the regression coefficients, the fully documented L-1.5 algorithm (Liem *et al.*, 2000) sufficed[17] for our present purposes in the absence of price variables. We recall its key features needed for the estimation of a U-GBC form, nested cases (R-GBC, TL, LL and LIN) of interest included, and essential for the production of derived statistics comparable across models, as these matter as well.

---

[12] For instance, if the true form is Trans-Log, Blackorby *et al.* (1977) and Denny & Fuss (1977) point out that output attribute functions embedded in the Trans-Log function must be log-linear. The former authors also remark that consistent aggregation also requires in that case that the output aggregator functions be Cobb-Douglas.

[13] Empirical credibility of the results naturally requires testing whether the data are compatible with the constraints or not, an often neglected test. For discussion of the constraints in the special TL case, see Denny & Pinto (1978).

[14] In practice, more general specifications allow the Box-Cox transformations for Quantities or Traffics to depend on the identity of the interaction terms on the right-hand side, which requires double indices, but this extension is unnecessary for present expository purposes.

[15] Cost share equations, which contain some of the regression parameters found in the TL form, are sometimes added to the cost level equation and jointly estimated in order to increase the precision of the TL parameter estimates: this may matter if samples are relatively small, as in Bereskin (2000) who had only 19 yearly observations (on 36 firms from 1978 to 1997) and needed to add firm and year dummies to the TL specification, making it less parsimonious still.

[16] The regularity conditions of the Shepard theorem (1953) for cost and production functions needed for the first derivative of the logarithm of the *level* of cost with respect to the logarithm of each input price to yield the cost minimizing cost *share* of each input, a property known as Shephard's Lemma, may not all hold for a monopoly firm where labour prices may not be exogenous and which may not always minimize total cost.

[17] In practice, we use the L-1.4 version (Liem *et al*., 1987) implemented in the TRIO software (Gaudry *et al*., 1993).

It is indeed important to note here that the user of BCT faces other barriers beyond the numerical one arising simply from the maximization of a multidimensional non linear function[18], barriers that may have prevented implementation of the GBC baseline. We must focus additionally on (i) the nature of the proper likelihood function to use for the U-GBC, (ii) *statistical* impediments due to special requirements (often ignored until 1984) on the proper calculation of *t*-statistics and due also to the necessary development of appropriate measures of the quality of adjustment (or "fit") based on expected values of the dependent variable; (iii) the burden of expressing[19] coefficient results in terms of *elasticities*. For our own results to be intelligible, some explicitation of these difficulties, not to say pitfalls, and of their resolution, matters as they jointly constitute a barrier.

The proper statistical model, an unweighted likelihood function? Simply stated, the classical Box-Cox regression model nesting all forms of interest for us is:

$$y_t^{(\lambda_y)} = \beta_0 + \sum_k \beta_k X_{kt}^{(\lambda_k)} + u_t \tag{1-A}$$

where, neglecting for simplicity dummy variables (not transformable by a BCT), the X_k denote any regression term, including interactions of variables, and both the dependent and the transformed independent variables are strictly positive. Note also that (1-A) contains the necessary intercept.

But this model presents a major challenge because the BCT on the dependent variable is defined only if $y_t > 0$: for this reason, there exist a number of statistical interpretations of (1-A), all coming to terms with this downward truncation of the explained variable in that model. Two of the main approaches are centred on Bickel & Doksum (1981), who modify the BCT power transformation to allow negative[20] values of y_t, and on a group of like-minded formulations by Spitzer (1976), Olsen (1978), Poirier (1978) and Amemiya & Powell (1980), who require that the distribution function of $y_t^{(\lambda)}$ contain no mass point. As the per kilometre maintenance cost can very well be extremely low, and even equal 0 for all practical purposes (for instance after reconstruction of the segment), the latter approach is not realistic for our problem. We use instead a third formulation extending the original single-limit Tobit model (Tobin, 1958).

As discussed at length in Dagenais *et al.* (1987), for problems like ours where y_t can be said to have both a lower limit (maintenance cost cannot be negative but can be close 0, implying the possibility of a mass point at $-1/\lambda$ when $\lambda > 0$, or in practice at a very small number ε) and an upper limit (maintenance cost cannot exceed the periodic reconstruction cost, called "renewal cost"), the proper statistical model to use is the Rosett & Nelson (1975) two-limit Tobit model where the exact likelihood of observing the vector $y \equiv (y_1, \ldots, y_t, \ldots, y_T)$ reduces, in the absence of limit observations in the sample, to the maximand Λ actually proposed by Box and Cox themselves:

$$\Lambda = \prod_{t=1}^{T} \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp\left(-\frac{u_t^2}{2\sigma_u^2}\right) \left|\frac{\partial u_t}{\partial y_t}\right| \tag{1-B}$$

[18] There are also additional *computational* loads to guarantee that a global maximum has been found because the likelihood function of y_t in (1-A)-(1-B) below is not necessarily unimodal.

[19] One could imagine software automatically calculating elasticities by numerical simulation, rather than analytically, but we have not found any.

[20] When the sample contains only positive observations, their Likelihood function is as in Box & Cox.

where, under the assumptions of normality and[21] constancy of the variance σ^2_u of the independent error term u_t of zero mean, the Jacobian of the transformation from u_t to the observed y_t is $\left|\partial u_t / \partial y_t\right| = y_t^{\lambda y-1}$. This is the rationale explicitly implemented in the L-1.5 algorithm which maximises the concentrated log-likelihood of (1-A)-(1-B) and provides derived statistics. As detailed values of all BCT are needed, it is not possible to dodge such maximization and simply find out whether one is closer to the Linear or the Log-Log case, as in Davidson & MacKinnon (1985).

One of these statistics is the probability, calculated for every observation of the sample, that it be a limit observation in the sense just mentioned[22] and that an unweighted likelihood function such as (1-B), apparently failing to recognize the truncation of the distribution of the u_t , be inappropriate for the data and specification at hand. The strict normalcy of the error distribution is therefore assumed to hold *ex ante* and its applicability tested *ex post* by ensuring that the actual sample contains no limit observations. The first and third right-hand side terms of Equation (8) below are estimated according to the procedure defined in Equations (56)-(57) of Liem *et al.* (1987). This strategy follows the remarks by Draper & Cox (1969) concerning the original BCT model, reiterated by Olsen (1978), that "the reasonableness of the [above] assumptions can always be tested by an analysis of residuals".

It seems more appropriate to recognize the presence of real lower and upper limits in (1-A) and to test afterwards for the probability of being too close to them than to make very strong *a priori* positive and unscrupulous assumptions about their absence. A balance can be found between stating that, for a model of say GNP, the Box-Cox maximand is not strictly a likelihood function because it fails to recognize that y_t is truncated (as GNP can be neither negative nor unduly high in a given sample) and the opposite stand consisting of always ignoring limits altogether.

Likelihood Ratio and Student-*t* tests. Concerning other derived statistics, Likelihood Ratio tests based on the χ^2 distribution are easily performed in the usual way, with standard information on the numbers of parameters estimated; but two Student-*t* tests are provided by the L-1.5 algorithm.

To understand the reason for this double till, an example will help. Tables of results such as those found in Appendix 3 indeed present two kinds of *t*-statistics: for the BCT, ***unconditional*** values calculated with respect to both 0 (the logarithmic case) and 1 (the linear case) and, for the β_k regression coefficients of transformed variables, *t*-statistics with respect to 0 that are ***conditional*** upon the estimated value of BCT form parameters.

The reason for this reporting asymmetry is that, in the presence of BCT, the unconditional *t*-values of β_k coefficients associated with transformed X_k depend on their "units of measurement" (Spitzer, 1984). An intuitive way to understand this result is to consider that BCT effect a reparametrization of the model whereby, among other things, if a particular β_k coefficient is assumed to be null, the regression constant cannot be assumed to remain unaffected, *ceteris paribus*, as in a standard linear regression conditions. But obtaining *t*-statistics[23] that depend on units of measurement of the transformed variables would make them useless because one could decide on desired values and then adjust the scale of all X_k accordingly: therefore, computing conditional *t*-values provides a useful[24] second best indicator for the β_k unless one wishes, as many now prefer, to rely exclusively

[21] The procedure allows, if necessary, for simultaneous corrections for heteroskedasticity and autocorrelation but they were not used for the models reported on here.
[22] And further explicated below in (8)-(9).
[23] This comes as a surprise to all of us who are used to the presumed invariance to units of measurement of *t*-statistics since their inception a century ago (Gosset, 1908).
[24] Consequently, all Box-Tidwell and Box-Cox results published before 1984, including all referenced above and below, as well as all computer programs currently available to estimate such models, have to be verified to make sure that their

on first best Log Likelihood ratio tests. These remain exact in the presence of BCT but are notoriously tedious to apply with large numbers of regressors[25] in the absence of an automatic procedure unfortunately not provided in the current L-1.5 algorithm version.

Fitted values and measures of overall adjustment quality. Another difficulty arising with model (1-A)-(1-B) pertains to measures of the quality of adjustment, or "fit", in the presence of Box-Cox transformations on the *dependent* variable. Consider both the adjustment with respect to individual observations and the overall measures of fit that are needed.

In evaluating the *goodness-of-fit with respect to each observation*, our interest:

(i) is, first and foremost, in how ***observed y***, not *transformed* y in (1-A), is accounted for. We cannot therefore use any statistic that simply relates calculated and transformed sample values of the regressand, as Ordinary Least Squared based programs applied to transformed variables automatically do. We must therefore focus on the "unrolled" value from (1-A):

$$y_t = f^{-1}[\ g(X_t) + u_t\ ]; \tag{2}$$

(ii) is also, secondly, in comparing the *observed y* with the *value produced by the full model*, which requires recognition of the fact that y is a random variable due to the presence of the random error. This demands that the *level of observed y* be compared to its *expected value*, not simply to a calculated value $y_t = f^{-1}[\ g(X_t)\ ]$ that would ignore the residual u_t and pretend that only the fixed (in a statistical sense) or deterministic part of (1-A) matters. And we will in any case presently note that such a calculated value is neither defined without problems nor separable and extractable from (1-A). The necessary calculation is therefore of:

$$E(\ y_t\ ) = E\{\ f^{-1}[\ g(X_t) + u_t\ ]\ \}\ \ . \tag{3}$$

The habit of neglecting the residual u_t comes of course from the Linear model where, as is well known, $E(y_t) = [\ g(X_t)\ ]$ because $E(u_t) = 0$ and the expected value expression (3) reduces to:

$$E(y_t \mid X) = \hat{y}_t\ \ . \tag{4}$$

But this convenient simplicity vanishes with Log-Log, Trans-Log and Box-Cox models. In the former two cases, the expected value of y derived from (3) is simply:

$$E(\ y_t\ ) = \{\ \exp[g(X_t)] * [\exp(u_t)]\ \} = k*\exp[g(X_t)] \tag{5}$$

where $\exp(u_t)$ is a log-normal random variable whose mean is a constant k over the sample and for which the maximum likelihood adjustment provides an estimate. In the more tricky Box-Cox case, we have, if λ denotes in (1-A) the BCT on the dependent variable:

$$E(\ y_t\ ) = E\{\ 1 + \lambda\ [\ g(X_t) + u_t\ ]\ \}^{1/\lambda} \tag{6}$$

---

t-statistics computed for β_k regression coefficients are conditional upon their estimated form parameters. Otherwise, as unconditional values, they depend on the scale of the transformed variables and are therefore arbitrary. Before 1984, many authors (*e.g*. Gaudry & Wills, 1978) estimated their models with grids for the BCT and, given estimated form values for the variables in (1), used OLS to derive the β_k and their conditional t-statistics: they were involuntarily in the right. Others estimated all model parameters jointly, without iterating between form and regression parameters, and calculated *t*-statistics unconditionally using first or second partial derivatives of the Log Likelihood. For instance, Berndt and Khaled (1979) themselves, already suspicious for using a cost share model without intercepts in the presence of a BCT (*op. cit*., Equation 28), further calculate unconditional *t*-statistics (*op. cit*., Table 4), making their *t*-tests on β_k coefficients unreliable as well.

[25] For instance, the Trans-Log model found in Column 3 of Annex 3 has 55 β_k coefficients.

where, by contrast with the previous case, the ratio between $E(y_t)$ and $f^{-1}[\ g(X_t) + u_t\ ]$ is not constant and needs to be calculated for each observation because the error term is unfortunately not separable from a calculated value term defined as:

$$y_t = \{\ 1 + \lambda\ [\ g(X_t)\ ]\ \}^{1/\lambda} \tag{7}$$

and the solution of which for y_t would raise problems of its own, notably with imaginary numbers, for certain values of λ. This problem would remain had one rejected the original Box-Cox model (1-A) in favour of an *ad hoc* approach consisting in adding an error term to (7) in order to separate the error. Due to these two impediments (solving the calculated value term for y_t and the impossibility of separating the error term from it in (6) as was done in the Log-normal case (5) above), it is best to avoid (7) and perform with the original Box-Cox model a proper calculation of the expected value in accordance with (3), as is generally, and unavoidably, done in other non linear models.

But we are not home free yet in calculating (6) because the fact that the domain of y_t is limited must still be accounted for in the determination of the expected value, even if the exact likelihood is (1.B): the calculation procedure must recognize both domain limits for y_t and Gaussian errors for u_t. Even in the absence of limit observations in the sample, if any observation y_t is assumed to be censored both downwards and upwards, $\varepsilon \le y_t \le \nu$, where ε and ν are respectively the strictly positive lower and upper censoring points common to all observations[26] in the sample, the generalization of the Tobit model to a doubly censored dependent variable formulated by Rosett and Nelson yields for the expected value of y_t defined in (3) the following expression:

$$E(y_t) = \varepsilon \int_{-\infty}^{u_t(\varepsilon)} \varphi(u)du + \int_{u_t(\varepsilon)}^{u_t(\nu)} y_t\varphi(u)du + \nu \int_{u_t(\nu)}^{\infty} \varphi(u)du \quad , \tag{8}$$

where $\phi(u)$ is again the normal density function of u with 0 mean and variance σ^2_u :

$$\varphi(u) = \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp\left(-\tfrac{u^2}{\sigma_u^2}\right). \tag{9}$$

Concerning now *overall goodness-of-fit*, the implemented algorithm provides two extensions of the usual single-equation R^2 measure of multiple correlation.

i) The first, the Pseudo-(E)-R^2, substitutes $E(y_t)$ from (8) for y_t in the well known and frequently used Pearson coefficient (*e.g.* Laferrière, 1999) consisting in the square power of the simple correlation coefficient between two variables, denoted Y_t and $E(Y_t)$ below. It is computed in unadjusted and adjusted (for degrees of freedom) variants, respectively (A) and (B), in:

$$Pseudo-(E)-R^2 \equiv \left\{ \begin{array}{lll} R_E^2 = \dfrac{\left[\sum_t^T (Y_t - \bar{Y})(E(Y_t) - \overline{E(Y)}) \right]^2}{\sum_t^T (Y_t - \bar{Y})^2 \sum_t (E(Y_t) - \overline{E(Y)})^2} \;, & & (A) \\ R_E^{*2} = \;\; 1 - \frac{T-1}{T-K}(1 - R_E^2)\;, & & (B) \end{array} \right\}, \tag{10}$$

[26] The original Tobin (1958) paper had limits that varied by observation.

where K is the total number of estimated parameters in (1-A) and T the number of observations. This R^2_E will always give values within the range [0 , 1] and be identical to the usual measure if the dependent variable y_t is linear.

ii) The second, the Pseudo-(L)-R^2 long ago specified by many authors (*e.g.* Aigner, 1971, pages 85-90) and based on logarithms of the likelihood function, is also computed in unadjusted and adjusted variants, respectively (A) and (B) in:

$$Pseudo-(L)-R^2 \equiv \begin{Bmatrix} R_L^2 = 1-\exp\{2(L_0 - L_{\max})/T\}, & (A) \\ R_L^{*2} = \quad 1-\frac{T-1}{T-K}(1-R_L^2)\ , & (B) \end{Bmatrix} \tag{11}$$

which is also within the [0 , 1] range since L_0, the sample maximum of the Log Likelihood associated with the constant β_0 as only regressor, is necessarily smaller than the maximum L_{max} obtained with more regressors. This R^2_L collapses to the previous measure R^2_E if (1-A) is linear.

Elasticities and marginal costs. Concerning difficulties specific to BCT, note finally that the β_k regression coefficients of variables transformed by a BCT have lost the intuitive contents they retained in simpler linear or logarithmic fixed forms. This increases the importance of elasticities as indicators of the reasonableness of results used in conjunction with indicators of the statistical significance of the β_k: significance never implies reasonableness.

For this reason, our detailed table of results found in Appendix 3 presents the classical (also called *sample*[27]) *elasticities* for all variables, including dummy variables[28]. Classical elasticities ignore the presence of the random term, treat y as deterministic, and are usually written:

$$\eta_t(y, X_k) = \frac{\partial y}{\partial X_k} \cdot \frac{X_k}{y} \Bigg|_{X_{kt}, y_t, X_{\ell t}} \tag{12}$$

where the vertical line simply means that the derivative is « evaluated at » $y = y_t$, at $X_k = X_{kt}$ for the variable of interest and at $X_\ell = X_{\ell t}$ for other right-hand-variables belonging to (1-A).

Expression (12) makes it clear that, even if (1-A) is linear, the value of this point sample measure of elasticity depends on the reference levels at that point and is therefore not constant across observations. The sample mean is used as reference point in Appendix 3:

$$\eta(\bar{y}, \bar{X}_k) = \eta_t(y, X_k) \Bigg|_{X_{kt} = \bar{X}_k,\ y_t = \bar{y},\ X_{\ell t} = \bar{X}_\ell} \tag{13}$$

[27] The expression “sample elasticity” is preferred to the less misleading “deterministic elasticity”. It can be evaluated anywhere, notably at the sample means, in which case it is the “sample elasticity evaluated at sample means”.

[28] In Appendix 3, code names of variables that are not underlined denote ordinary strictly positive variables and code names that are underlined twice denote dummy variables. Such underlines imply distinct formulas to compute the associated elasticity: one for ordinary strictly positive variables and one for dummy variables. We will not provide here the analytical expressions for the “elasticities” calculated for the dummy variables, describing in percentage terms the impact of the presence of the dummy, but the documentation of the algorithm referenced above is easily consulted. In a few words, they imply multiplying the standard formulas documented here by the ratio of the mean of the dummy variable computed over its positive values (typically 1) to the sample mean of the dummy variable (zeroes included). The adequacy of this approximation of the discrete jump in E(y) implied by the presence of the dummy variable, as compared with its absence, is discussed at length in Dagenais *et al.* (1987) where other measures of “percentage changes of y or E(y) due to the presence of the dummy variable” are also provided.

In the TRIO program L-1.4 algorithm, the user can also request the *Expected value elasticity*:

$$\eta_t\,(E(y),\, X_k) \;= \frac{\partial\, E(y)}{\partial\, X_k} \cdot \frac{X_k}{E(y)} \Bigg|_{X_{kt},\; E(y_t),\; X_{\ell t}} , \tag{14}$$

where the partial derivative of (8) is required and evaluation is effected at the sample means. However, in tables and graphs forthcoming below, the elasticity of $E(y_t)$ with respect to traffic, and the marginal (Expected) cost obtained by multiplying it by $E(y_t)/X_{kt}$, will be calculated as means of individual values obtained for each observation according to an approximation[29] of (14), rather than at the means of the sample. The *mean elasticity* $\bar{\eta}^*$ so obtained from approximations at all sample points, and the *mean marginal cost* $\bar{\mu}^*$ derived from it, are then simply, in turn:

$$\bar{\eta}^*\,(E(y), X_k) \;= \sum_{t=1}^{T} \eta_t\,(E(y), X_k) \Big/ T \tag{15}$$

and

$$\bar{\mu}^*(E(y), X_k) \;= \sum_{t=1}^{T} \left[\eta_t\,(E(y), X_k)\right] \bullet \left[E(y)_t\,/\,X_{kt}\right] \Big/ T\,. \tag{16}$$

A feasible following for Khaled's 1978 innovation? These remedies show that the "needless complications" evoked by Caves *et al.*, (1980a) for proper Box-Cox models and the derived "interpretative difficulties" of BCT pointed to in high level surveys (*e.g.* Jara-Diaz, 1982; Tovar *et al.*, 2003) exist but constitute insufficient barriers to prevent the seminal Berndt & Khaled (1979) baseline demonstration from having a following.

The remedies of interest here make it possible to recognize that track cost measured on small network sections differ from firm-wide total costs in that they are relatively small and have non negligible chances of including limit observations, a matter of smaller concern if one is dealing with total U.S. manufacturing output or even with the total yearly transport output of very large American or Canadian rail companies. They also make it possible to use proper common statistics to compare the competing forms.

An important additional barrier, related to the treatment of zeros in variables subjected to a BCT, remains to be addressed in the next section but will not invalidate this approach.

[29] There are therefore two differences between the cost elasticities with respect to traffic listed in Appendix 3, calculated according to (13), and those used in forthcoming graphs and tables, calculated by (15), an approximation of (14). The first difference is between an evaluation at sample means and the mean of evaluations at each sample point. The second is that in (15) we use an approximate expected value measure, not an exact one. What does the approximation consist of? We did not calculate $E(y_t)$ for each observation exactly according to (14) but approximated it by multiplying the "unrolled" calculated value term (7) by the average ratio (across all observations) between (6) and (7) and not by the observation-specific ratio. As the latter actually differs across observations, our time saving approximation assumes that the difference is stable across observations for a given regression specification and stable across regression variants whose elasticities (or derived marginal costs) are to be compared. We performed numerous comparisons of (7) and (6) vectors to verify that this is in fact a weak assumption in our problem.

4. Design of form tests and selection of maintained hypotheses from exploratory work

Developments above show that it is possible to develop a common statistical framework built on a maximum likelihood estimator and its derived results, such as elasticities and measures of goodness-of-fit, in order to compare the U-GBC to simpler nested forms. Before proceeding to full trials in Section 5, including tests pertaining to the level of output aggregation, competing forms must be specified; in addition, we take advantage of previous exploratory work to formulate "maintained hypotheses" as guides to the by-the-book formal sequel of Sections 5 and 6.

4.1. The neglected problem of zero values and the selection of competing forms

Abstracting for the moment from the question of the number of train traffics (the level of output aggregation), the structure of forthcoming tests, found in Table 4, respects two sets of constraints.

First, forms matter more for some groups of variables than others: for instance, as State variables S consist almost exclusively in dummy variables, we prefer to ignore their interactions with other variables. This limits in practice to 6 the number of candidate groups of variables for the general U-GBC form: we designate them as S, Q, T, Q*Q, Q*T and T*T, where single letters denote vectors of appropriate length and stars denote products of 4x1 vectors to be fleshed out in Table 9. The 4 forms nested into U-GBC and mentioned earlier are those defined on lines 1-4 of Table 4.

Second, it is not possible to select the forms to be compared without simultaneously raising the issue of zeroes to ensure that chosen candidate specifications are all invariant to changes in "units of measurement", strictly understood as *scale changes*, of the transformed variables. This second requirement, applied in the Z-0 to Z-3 columns of Table 4, requires further explicitation. Only alluded to above in a footnote to our presentation of the Generalized Translog Multiproduct Cost Function (GTMCF), the requirement has yet received insufficient attention in the literature but is now addressed as a systematic dimension of the design of detailed specifications to be tested.

Table 4. Feasibility of competing forms according to the rule adopted for the transformation of zeroes

		Simplified model form specifications[A]	Feasibility of the form under			
	Form of	**C = f(S , Q , T , Q*Q , Q*T , T*T)**	**Z-0**	**Z-1**	**Z-2**	**Z-3**
1	***Linear***	C = lin [S ; Q ; T]	Yes	Yes	Yes	Yes
2	***Log-Log***	ln C = ln [S ; Q ; T]	***No***	Yes	Yes	Yes
3	***Trans-Log***	ln C = ln [S ; Q ; T ; Q*Q ; [T*Q] ; (T*T)]	***No***	Yes	***No***[B]	Yes
4	***R-GBC***	$C^{(\lambda)}$ = bc [S ; $Q^{(\lambda)}$; $T^{(\lambda)}$; ---- ; $[T*Q]^{(\lambda)}$; $(T*T)^{(\lambda)}$]	***No***	Yes	Yes	Yes
5	***U-GBC***	$C^{(\lambda y)}$ = bc [S ; $Q^{(\lambda x1)}$; $T^{(\lambda x2)}$; ---- ; $[T*Q]^{(\lambda xk)}$; $(T*T)^{(\lambda xK)}$]	***No***	Yes	Yes	Yes

[A] On lines 3-5, the few continuous variables found in **S** are transformed as required by the definition of each form and the interaction terms written as **X*X** include only off-diagonal terms $\mathbf{X_i X_j}$ **(i ≠ j)** because diagonal terms $\mathbf{X_i X_j}$ **(i = j)** are subsumed[30] in the $\mathbf{X_i}$ terms due to the fact that $\mathbf{[\,X_i^{(\lambda ii)}\, X_i^{(\lambda ii)}\,] = [\,X_i^{2}\,]^{(\lambda ii)} = [\,X_i^{(\lambda i)}\,] = [\,X_i^{N}\,]^{(\lambda ii)}}$. On Line 5, the four BCT indices of the U-GBC do not denote the exact number of distinct BCT actually used (as many as 10 in cases reported here, and more in some unreported cases) but simply express the generality of the U-RBC, as compared to its constrained single-BCT origin.

[B] Impossibility due some interaction terms that cannot be defined because they contain ratios of zeroes. In that case, an associated dummy variable cannot be a remedy, as it can be in a simpler case containing only products of zeroes.

Invariance to scale of measurement when strictly positive variables are transformed. The double-limit "Box-Cox Tobit" statistical framework documented in the previous section noted the importance of the *invariance to units of measurement* (of transformed strictly positive dependent

[30] The proof that the Box-Cox transformation is invariant to a power transformation of the transformed variable even in the absence of a regression intercept can be found in Gaudry & Laferrière (1989).

and independent variables) of two critical sets of estimates: (i) the *BCT form estimates* themselves, which impose a regression intercept (Schlesselman, 1971); (ii) the estimates of the *t-statistics of* β_k *regression coefficients* which, in contrast to those of the form parameters, need to be computed conditionally upon (duly invariant) values of the BCT form estimates (Spitzer, 1984). We now focus on a third invariance pitfall arising with transformed variables containing zero values.

In our application, transformed traffic output levels are sometimes null for some types of trains, a quite frequent situation in other contexts with some particular explanatory variables[31] (*e.g.* snowfall). The issue has long been recognized, for instance in Winston (1985, page 63):

> "*To be sure, the Trans-Log approximation runs into difficulty for zero values of output. In this case, a transformation using the Box-Cox metric (Caves, Christensen & Tretheway, 1980a) can be used to apply this functional form.*"

If it is obvious that zero values are not easily handled when a logarithmic transformation of a variable is required, it is less obvious that, contrary to what the quotation implies, the same problem can arise in models making use of BCT where, depending on how it is handled, it might compromise form estimates and invalidate the desired comparisons among competing model forms.

Invariance to scale of measurement when transforming variables that contain zeroes. In order to avoid this new pitfall, we distinguish in Table 4 between four ways, or Z-rules, of handling a variable X_k containing observed zero values, namely:

- Z-0. Transform all observations on the variable, zero values included, by a BCT;
- Z-1. Replace the zero values by a very small value before applying a BCT to the variable;
- Z-2. Add an associated dummy variable to a transformed variable that contains some zero values;
- Z-3. Remove all observations containing zero values before applying a BCT to the variable;

only two of which (Z-1 and Z-2) provide an invariant BCT form estimate, short of avoiding the difficulty altogether with (Z-3). We comment on the first three and will apply the last three below.

The Z-0 "*ignore the difficulty*" rule[32] fails to "compensate" for the $-1/\lambda$ shift at zero values resulting from the application of the transformation to these values[33]. This rule unfortunately produces BCT estimates that depend on the scale units of the variable, a pitfall demonstrable by a simple numerical exercise on the computer. All form estimates resulting from the application of Z-0 are then dubious, as are also any unconditional *t*-statistics derived from matrices of partial derivatives of the Log Likelihood with respect to all parameters including the scale dependent BCT estimates. Curvature of a cost function should not depend on arbitrary units of measurement.

The zero *replacement* rule Z-1, an approximate approach, is a satisfactory option adopted in most of our tests. We will show that the implications of this rule vary to some extent across models forms but have more impact on LL and TL than on R-GBC and U-GBC forms. This may be due to the fact

[31] The possibility of zero values of the dependent variables constitute a distinct issue that we will not address beyond what was summarily said (the discussion of the full issue takes up many pages in the published references) in the discussion of the calculation of the expected value of y with equation (8).

[32] It was applied in the GTMCF single-BCT Box-Tidwell tests carried out by Caves *et al.* (1980a, 1980b and 1985) mentioned earlier and referred to by Winston. From what we could determine, the estimated value was found by a non linear search without using analytical first derivatives: as the derivative of a variable V transformed by BCT with respect to λ involves the natural logarithm of the observations (it is equal to $\{[1/\lambda]\bullet[V^{\lambda}\ln(V) - V^{(\lambda)}]\}$ if $\lambda \neq 0$ and to $\{(1/2)\ln^2(V)\}$ if $\lambda = 0$), the maximization would have aborted if analytical derivatives had been used.

[33] For a positive BCT, the transformation is defined at zero output as $[\ (0^{\lambda}-1)/\lambda = -1/\lambda\ ]$. But this shift of $-1/\lambda$ is clearly unaffected by a scalar transformation of the X_k vector of values, hence the lack of invariance: being well defined does not imply being invariant to scale changes.

that, under Z-1, the logarithmic transformation generates very large negative values that can behave as outliers, a bias that is avoided when the optimal BCT differs from zero, as it typically does below. It would not be a satisfactory option for LL and TL forms if, contrary to our case, there were too many zeroes in the sample.

The "*add a dummy*" rule Z-2, the strict approach, is used in a sub-section below where the influence of zero values is specifically explored: it is a tedious but exact remedy in all Log-Log, Trans-Log and Box-Cox models (but, as indicated in Table 4, not in the special case of our Trans-Log models as specified because they involve ratios of zero values). It requires, for each standard variable that includes such values (and may conveniently be called for this reason a *quasi-dummy variable*), building an *associated dummy variable* that compensates for the missing shift at 0 when the scale of the variable transformed by a BCT[34] is modified.

The L-1.4 algorithm[35] that we used (Liem *et al.*, 1993) generates such additional dummy regression variables automatically and defines them conventionally as equal to 1 if the quasi-dummy variable is positive[36] and to 0 otherwise. Naturally, only one such associated dummy is admissible if two or more transformed variables have zero values in the same locations, lest one wished to live with exact collinearity! In our most general U-GBC model, only 4 associated dummies will be in fact be needed despite the much larger number of quasi-dummy variables transformed.

In addition to this scale invariance preservation role, an associated dummy variable provides an unexpected benefit, documented in Appendix 4, because it amounts to a decomposition of the effect of the quasi-dummy variable between a qualitative (*e.g.* the effect of the presence of snow) and a quantitative (*e.g.* the effect of the amount of snowfall) component: this decomposition is a substantive issue of general relevance in regression analysis involving quasi-dummy variables.

4.2. Maintained hypotheses from explorations on LARGE database with pared down models

The test structure design found in Table 4 represents a tall order, especially if aggregation of traffic classes is also considered. It is therefore both relevant and useful to report on exploratory work carried out with pared down 5-traffic model specifications because such tests based on minimal sets of explanatory variables lead to the formulation of "maintained hypotheses" for further use.

We therefore first summarize the main results[37] drawn from the exploratory version of this paper (Gaudry & Quinet, 2003) where a first analysis was made using the LARGE database under rule Z-1 after the following choices were imposed on the possibilities implied by Table 1:

---

[34] In practice, the zero values are then left untransformed in our Log-Log and Box-Cox models, both of which having associated dummy variables added: the procedure is documented in Appendix 4. Strictly speaking one should distinguish between changing the scale and changing the units of measurement of a variable, but they are here confounded because "changes in units", as commonly designated loosely, are assumed to simply modify the scale of variables: clearly, it is easy to imagine changes in units of measurement that are not equivalent to scale changes, *e.g.* transforming Fahrenheit into Celsius degrees.

[35] The three invariance pitfalls are also found in Box-Cox Logit estimation. For instance, the very recent BIOGEME software (Bierlaire, 2003, 2008) calculates unconditional *t*-statistics for all parameters of this model, β_k included. In this case, an additional iteration should be effected conditionally upon the estimated BCT values in order to obtain conditional t-statistics of the regression coefficients.

[36] But using the complementary definition would simply change the regression sign of the associated dummy variable coefficient and have no effect on other parameters and on the Log Likelihood.

[37] This widely circulated paper, available from a number of web sites indicated in the reference, gives the full set of final regressions on all forms but the summary presented here is self-standing.

i) **Variables:** the only ***Q*** variable included was the *maximum allowed speed*, assumed to be a good indicator of the standing of a line, but ***five T*** variables were included, thereby treating the (input) track *maintenance trains* on the same basis as the passenger and freight (output) services;

ii) **Models:** only 12 explanatory variables were retained for each form, in the hope that selecting only the most significant ones would give fair clues to the nature of the form problem. The models were effectively pared on statistical grounds, without due consideration for economic theory or engineering practice.

U-GBC form dominance. The results in Table 5 show a clear pattern: the LL and TL models are roughly equivalent, but the LIN model is massively inferior, and the GBC models are massively superior, to them.

Table 5. Exploratory One-Quality, Five-Traffic, Twelve-Variable models (1146 observations)

Model	A	B	C	D	E
Specification of form	**LIN**	**LL**	**TL**	**R-GBC**	**U-GBC**
Log likelihood LL	-16 069.7	-15533.9	-15 535.8	-15 359.7	-15 322.9
Beta regression coefficients	12	12	12	12	12
Box-Cox transformations	0	0	0	1	7
Cost elasticity with respect to traffic (*)	0,42	0,07	0,34	0,36	0,36

(*) Calculated with respect to weighted traffics and evaluated at the sample means of observations.

As one expects significant non linearity, the surprise is not the poor performance of the Linear form but that of the current workhorse of maintenance cost analysis, the Trans-Log. As the Trans-Log specification retained was not kept whole in the paring down exercise, it became almost indistinguishable from a Log-Log (CES) model. Retaining only the most significant terms of the Trans-Log form biases the comparison in its favour here, as it then has about the same number of estimated parameters as the other forms. Despite this generosity, the 1-parameter R-GBC dominates it by 176 Log likelihood points (with one degree of freedom of difference) and the U-GBC, allowing for a specific BCT on the dependent variable C and for 5 additional train-specific BCT, raises the advantage to 213 points (with a difference of seven degrees of freedom). It remains to be seen in Section 5 whether a less Procrustean and less parsimonious procedure yields otherwise.

Identifying maintenance requirement power functions by traffic class. This last difference of 36.8 Log Likelihood points between the R-GBC and the U-GBC pointed to the need for more details on the BCT. In consequence, the path of generalization from the former to the latter form is made explicit in Table 6 where, from model B to model G, each train output measure successively obtains its own BCT. As the Log Likelihood increases from 15349.0 to -15322.9, a gain of 26 points (for 5 degrees of freedom), the evidence suggests that each train type generates specific maintenance costs. All train-specific BCT are positive and between 0,46 and 1,08.

Further refinements. Further refinements of model G carried out involved: (i) searching for various specifications of potential interactions among Traffic types, (ii) including regional dummy variables to account for systematic cost differences among 19 administrative divisions (out of 23) of the French National Railways (SNCF) performing track maintenance represented in the sub-sample; (iii) distinguishing between the number of trains by category and their average weight instead of using gross tonnage by train type as the relevant Traffic output measurement variable.

Generally speaking, the most sophisticated refinements won easily in terms of statistical fit, although this was somewhat less decisive for (iii), but no trial modified the general bearing of the results of Table 5 and Table 6, at least in statistical terms. Derived results from (iii), based on a specification containing 9 BCT (2 more than the number used in Column G of Table 6 due to the break-up of Traffic indicators between numbers of trains and their average weight), were: (a) that the elasticity of cost with respect to traffic level (evaluated at the mean value of the observations) was equal to 0,37; (b) equivalence coefficients between types of traffics were as shown in Table 7.

Table 6. From R-GBC to U-GBC by successive addition of specific BCT to traffic classes

<table>
<tr><th></th><th>Model</th><th>A</th><th>B</th><th>C</th><th>D</th><th>E</th><th>F</th><th>G</th></tr>
<tr><th>Variables</th><th colspan="8">Box-Cox transformations and their unconditional t-statistics with respect to 0 and 1</th></tr>
<tr><td rowspan="3">Cost per km
(C)</td><td>λ estimate</td><td>0.255</td><td>0.240</td><td>0.240</td><td>0.243</td><td>0.245</td><td>0.246</td><td>0.247</td></tr>
<tr><td>t-stat. w.r.t. 0</td><td>[26.69]</td><td>[25.96]</td><td>[25.47]</td><td>[26.00]</td><td>[26.85]</td><td>[25.91]</td><td>[26.53]</td></tr>
<tr><td>t-stat. w.r.t. 1</td><td>[-77.86]</td><td>[-82.01]</td><td>[-80.73]</td><td>[-81.00]</td><td>[-82.55]</td><td>[-79.38]</td><td>[-80.66]</td></tr>
<tr><td rowspan="3">Intercity trains
(GL)</td><td>λ estimate</td><td rowspan="15">As above and below</td><td rowspan="15">As below</td><td>0.617</td><td>0.520</td><td>0.546</td><td>0.530</td><td>0.469</td></tr>
<tr><td>t-stat. w.r.t. 0</td><td>[5.12]</td><td>[4.93]</td><td>[4.96]</td><td>[4.82]</td><td>[4.34]</td></tr>
<tr><td>t-stat. w.r.t. 1</td><td>[-3.18]</td><td>[-4.55]</td><td>[-4.12]</td><td>[-4.28]</td><td>[-4.92]</td></tr>
<tr><td rowspan="3">Regional trains
(TER)</td><td>λ estimate</td><td rowspan="12">As below</td><td>1.074</td><td>1.114</td><td>1.059</td><td>1.098</td></tr>
<tr><td>t-stat. w.r.t. 0</td><td>[5.28]</td><td>[5.51]</td><td>[5.21]</td><td>[5.31]</td></tr>
<tr><td>t-stat. w.r.t. 1</td><td>[0.36]</td><td>[0.57]</td><td>[0.29]</td><td>[0.47]</td></tr>
<tr><td rowspan="3">I-de-France trains
(IdF)</td><td>λ estimate</td><td rowspan="9">As below</td><td>0.735</td><td>0.742</td><td>0.706</td></tr>
<tr><td>t-stat. w.r.t. 0</td><td>[5.87]</td><td>[5.79]</td><td>[5.67]</td></tr>
<tr><td>t-stat. w.r.t. 1</td><td>[-2.12]</td><td>[-2.01]</td><td>[-2.36]</td></tr>
<tr><td rowspan="3">Freight trains
(F)</td><td>λ estimate</td><td rowspan="6">As below</td><td>1.348</td><td>0.979</td></tr>
<tr><td>t-stat. w.r.t. 0</td><td>[2.46]</td><td>[1.91]</td></tr>
<tr><td>t-stat. w.r.t. 1</td><td>[0.64]</td><td>[0.04]</td></tr>
<tr><td rowspan="3">Maintenance trains
(HLP)</td><td>λ estimate</td><td rowspan="3">As below</td><td>0.813</td></tr>
<tr><td>t-stat. w.r.t. 0</td><td>[1.47</td></tr>
<tr><td>t-stat. w.r.t. 1</td><td>[-0.34]</td></tr>
<tr><td rowspan="3">All X_k variables, including traffics without specific λ</td><td>λ estimate</td><td>0.255</td><td>0.430</td><td>0.372</td><td>0.325</td><td>0.213</td><td>0.138</td><td>0.122</td></tr>
<tr><td>t-stat. w.r.t. 0</td><td>[26.69]</td><td>[10.66]</td><td>[7.75]</td><td>[7.15]</td><td>[4.78]</td><td>[2.91]</td><td>[2.24]</td></tr>
<tr><td>t-stat. w.r.t. 1</td><td>[-77.86]</td><td>[-14.15]</td><td>[-13.11]</td><td>[-14.84]</td><td>[-17.66]</td><td>[-18.10]</td><td>[-16.13]</td></tr>
<tr><th colspan="9">General statistics</th></tr>
<tr><td colspan="2">Log Likelihood value</td><td>-15359.7</td><td>-15349.0</td><td>-15347.1</td><td>-15339.7</td><td>-15330.4</td><td>-15323.8</td><td>-15322.9</td></tr>
<tr><td colspan="2">Number of estimated β_k coefficients</td><td>12</td><td>12</td><td>12</td><td>12</td><td>12</td><td>12</td><td>12</td></tr>
<tr><td colspan="2">Number of estimated λ powers</td><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr>
</table>

Table 7. Relative marginal cost and virtual revenue from marginal cost charges by user class

Type of traffic	Long distance passenger	Regional passenger	Ile-de-France passenger	Freight
Relative marginal cost equivalence coefficient	1	5,5	1,9	0,4
Cost recovery rate under marginal cost pricing	0,07	0,11	0,15	0,06

Maintained hypotheses for more thorough trials. The main criticism made of this exploratory work was that the paring down had been arbitrary and too exclusively done on statistical grounds, an approach that made the proper detection of interactions among variables difficult. It was also stated that the presence of track maintenance trains, an input to production, in an equation used to derive the marginal costs of 4 user services, was questionable. Finally, the fact that *maximum allowed speed* was very statistically significant argued for new tests making full use of the 3 other Quality factors present in the database, even at the cost of a moderate reduction in the number of available observations (from 1146 to 985 observations).

Overall, these results suggest maintained hypotheses to be tested in the extensive trials:

(a) the U-GBC form is overwhelmingly dominant (Table 5); (H_0-1)

(b) axle-load power damage functions, identifiable only if traffic is broken down by class, vary by traffic class (Table 6); (H_0-2)

(c) distinct user class damage functions imply different marginal costs and recovery rates from a potential marginal cost pricing *régime* (Table 7). (H_0-3)

5. Form, aggregation and zero-output values: a preferred Four-Traffic U-GBC model

5.1. Maximal detailed contents: a Four-Quality, Four-Traffic, Full Trans-Log model

The groups of variables (S, Q, T, Q*Q, Q*T and T*T) included in the five competing forms of Table 4 and their detailed contents in terms of variables must be determined. We aimed first at having as a reference the richest Trans-Log specification possible in terms of groups and their detailed contents and then at defining other forms by nesting restrictions on the presence of the reference TL groups, as follows:

i) ***Variables***

- **S**: use all 11 available State variables, all Boolean except for *Number of track apparatus* and *Length of section*, both of which are transformed in accordance with the form requirement;
- **Q**: use all 4 available Quality factors;
- **T**: use the 4 traffics that correspond to outputs and neglect the 5th, an input factor;

ii) ***Interactions***[38]

- **Q*Q**: keep interactions among qualities **Q** in the Trans-Log reference form but remove them from the Box-Cox forms because, in background tests, they add nothing within these models. This will slightly affect the strict nesting of the Trans-Log form in the Box-Cox forms: if necessary, use non-nested tests to compare such models;
- **Q*T**: define interactions among track qualities **Q** and traffics **T** (i) as products of *Age of rails* or *Age of sleepers* (traverses) and traffic levels; and (ii) as ratios of traffic levels and *Maximum allowed speed* or *Proportion of concrete sleepers*. The reasons for defining ratios, as opposed to products, are discussed in Appendix 1. Note in passing that the ratio form of interaction is only identifiable with Linear and Box-Cox forms: in Log-Log and Trans-Log forms, ratios cannot be identified and distinguished from products;
- **T*T**: use all interactions among traffics **T**.

Contents being clarified, Table 8 presents results according to three dimensions: functional form, as defined in Table 4; level of traffic aggregation; rules adopted for the handling of zero values.

Comments in the forthcoming sub-section 5.2 (a) establish our 4-traffic preferred model on the basis of an examination of the doubly framed first line values and of the triply framed last column of Table 8; (b) probe the advantages of our chosen model in comparison with all other non linear results in Table 8. Sub-section 5.3 explores the role of interaction terms across TL and GBC models to understand the reasons for the better performance of the GBC models. Sub-section 5.4 reconciles our implied cost recovery results to other European results.

5.2. Form, aggregation and the key role of zero handling rules

A. *Embarras de richesse* : a tridimensional table

Two sets[39] of parallel test were carried out to establish the reference model series indicated in Table 8: the first one, with 985 observations, included the 18 high speed only links and the second one, with 967 observations, did not. As both series gave extremely close results, only those obtained with the MEDIUM database of 967 observations are shown in Part I of Table 8 where rule Z-1 is applied, as it was for the exploratory model of Table 5. In Part I, one also finds results obtained with only 2 traffics (passenger and freight) as well as results obtained with a 1-traffic aggregate. Part II

[38] In Section 4.1, we have already excluded interactions among state variables S and all other variables.

[39] The results are described in full in Gaudry & Quinet (2009). It was decided not to retain the 18 high speed links, insufficiently numerous to be treated separately and insufficiently homogeneous with the rest of the links in terms of technical characteristics and traffic mix, thereby probably reducing the precision of overall estimates.

presents some of the same models under rule Z-2, as does Part III for estimates under rule Z-3 with the SMALL database.

The table dimensions (namely *form*, *aggregation* and the *handling of zeroes*), impose some discipline on our comments. We will start with a focus on the first line and last column to establish our choice of a preferred model. Afterwards, we will extend the comments to the rest of the table results. Note that, in Table 8, all models are labelled "A" to mean that they are specified with the maximum number of variables just allowed above for the 5 forms in Table 4. Other less complete models will be examined later and their corresponding series of results labelled "B" and "C".

Table 8. Core model Log Likelihood values under varying form, aggregation and handling of zeroes

Model	**A**	**B**	**C**	**D**	**E**
Form as specified in Table 4	**LIN**	**LL**	**TL**	**R-GBC**	**U-GBC**
Part I. Zero *replacement* rule Z-1 (967 observations)					
I.4 Series A: [4 Traffics]		***Reference model series***			***Preferred***
Log-Likelihood	-13 547,642	-13 129,267	-13 092,508	-12906,375	**-12 866,366**
Number of parameters estimated	***19***	***19***	***55***	***42***	***51***
Difference w.r.t. Log	0	0	36	23	32
I.2 Series A: [2 Traffics]					
Log-Likelihood	-13 576, 284	-13 137,707	-13 099,112		-12903,667
Number of parameters estimated	***17***	***17***	***38***		***35***
Difference w.r.t. Log	0	0	21		18
I.1 Series A: [1 Traffic]					
Log-Likelihood	-13 600,709	-13 118,611	-13 097,894		-12 921,412
Number of parameters estimated	***16***	***16***	***31***		***29***
Difference w.r.t. Log	0	0	15		13
Part II. Zero *maintenance* rule Z-2 with associated dummy variables (967 observations)					
II.4 Series A: [4 Traffics]					
Log-Likelihood	-13 547,642	-13 119, 643	n.a.		-12849,541
Difference w.r.t. Log	- 4	0	n.a.		37
II.2 Series A: [2 Traffics]					
Log-Likelihood	-13 576, 284	-13 122,034	n.a.		-12 900,967
Difference w.r.t. Log	-2	0	n.a.		19
II.1 Series A: [1 Traffic]					
Log-Likelihood	Same as in part I				
Difference w.r.t. Log					
Part III. Zero *removal* rule Z-3 (928 observations)					
III.4 Series A: [4 Traffics]					
Log-Likelihood	Not feasible reliably: only 208 observations				
Difference w.r.t. Log					
III. 2 Series A: [2 Traffics]					
Log-Likelihood	-13 037,526	-12617,078	-12593,371	-12420,669	-12 396, 476
Difference w.r.t. Log	0	0	21	10	18
III. 1 Series A: [1 Traffic]					
Log-Likelihood		-12610,610	-12592,016	-12421, 338	
Difference w.r.t. Log		0	15	6	
Note: series A contains the maximum number of variables in each form, in accordance with Table 4. Series B and C containing reduced numbers of variables will be presented in Table 12.					

B. A statistically preferred 4-traffic U-GBC model: the rows and the last column of Table 8

Looking from left to right: an optimal form in the last column. The pattern seen in Table 5, symbolically expressed as [LIN] < [LL ≡ TL] < [R-GBC] < [U-GBC], is present in all three parts of Table 8 and on all lines. For instance, the pattern arising from the comparison of doubly framed

Log-likelihood values of the first line 4-traffic model under Z-1 is representative of what is found on other lines and can be described in the following words:

(i) the Linear form is dominated by all non linear forms in terms of Log Likelihood. As shown in Appendix 3 (Part III, Line 3), it is also the only form with a positive probability of limit observations: in all non linear cases, the use of Likelihood function (1-B) is justified *ex post*;

(ii) the Trans-Log model form is roughly equivalent to the Log-Log model because 36 degrees of freedom are involved in the comparison, for a gain of only 36,375 Log Likelihood points;

(iii) the U-GBC form dominates all. It is out of reach of the models based on logarithms (Log-Log and Trans-Log), both of which perform much less well, and even of the R-GBC[40] because the additional 9 BCT allow for a gain of some 40 Log Likelihood points, *i.e.* well above the 99% significance level of the $\chi^2/2$ distribution.

As this pattern, consistent with (H_0-1), is repeated on all lines[41] independently from the degree of aggregation or from the way in which the zero values are handled, the selection of a preferred model must involve a choice among the (triply framed) Log-Likelihood values found in Column E.

Looking up and down the last column: optimal aggregation on the first line. Choosing among Column E outcomes to determine whether (H_0-2) is also supported involves the other dimensions.

Concerning aggregation, simple non-nested tests are easy to perform within each part of the table because the number of observations is constant for all tests performed under a given Z-rule. Such tests involve the construction of hypothetical models made up of the union of the sets of all regressors, say X_I and X_{II}, found in the models of the pair considered (I and II) where the individual members differ only by the traffic aggregates. Consequently, the imposition of 0 restrictions on either subset yields the other member of the pair: such nesting amounts to a test of the level of aggregation because the restrictions pertain only to the variables representing traffic aggregates.

Performing this exercise for each *given Z-rule* among the Column E results of Table 8 supports the most disaggregate 4-traffic model, as opposed to more aggregate models. However, before taking this as a clear establishment of the U-GBC model of line 1 and as a confirmation of (H_0-2), we must be sure that an examination *across Z-rules* also validates the same choice.

To probe this important issue further, we consider in Table 9 more detailed information on all models from Column E. This table presents only the BCT values of (first order) traffic variables and their indicators of reliability under the different aggregation and zero value handling conditions of Column E: the other parameters are neglected. Many points pertain directly to (H_0-2):

i) **Aggregation.** As number of traffics decreases from 4 to 1, holding the Z-rule constant, the adjustment is not the only worsening indicator: the range of estimated traffic power values also decreases along with their reliability, as indicated by collapsing (unconditional) *t*-statistics[42]. At the 1-traffic limit of aggregation, the power estimate of 0,779 differs neither from 0 nor from 1:

[40] A number of other studies (Andersson, 2009a, 2009b; Marti *et al.*, 2009; Link, 2009) have found Box-Cox transformations applied to own terms to dominate the Log-Log forms, but as none have used full R-GBC specifications with interaction variables, strict comparisons are difficult to make.

[41] Due account taken of missing cases. Naturally, the Log-Likelihood gains differ within each zero handling regime and according to the level of aggregation. For instance, as one aggregates from 4 to 1 traffic under Z-1, the number of points gained per degree of freedom sacrificed is consistently low [1,02; 1,84 and 1,38] if one compares the TL with the LL and fall gradually [8,22; 13,00 and 15,17] if one compares the U-GBC with the LL, but these fluctuations do not argue against our preferred model choice.

[42] Correctly interpreted by keeping in mind that one is calculated with respect to 0,000 and the other with respect to 1,000 using the formula $t = [(\beta_h - \beta_0)/ \sigma(\beta_h)]$.

one can reject neither the logarithmic nor the linear forms ─anything will do! Aggregation clearly kills the power issue;

ii) **Z-rules.** The impact of using Z-2 instead of Z-1 for the preferred model is related to the proportion of zero values in the transformed variables: little effect on the BCT of the Freight traffic variable (no zeroes), a small impact on the Intercity traffic variable (1% of zeroes) and a very large impact on the Regional train variable (36% of zeroes for both types of trains considered together), as is indicated in the doubly framed cells. This exact Z-2 remedy requires 9 associated dummy variables but the resulting increase in the number of parameters causes a gain of only 16,92 Log Likelihood points and poorer *t*-values with respect to both 0 and to 1, thereby indicating a much greater uncertainty[43] of estimates than under Z-1;

Z-1 not a bad compromise. If the removal of all zero replacements (here equal to 0,00001) yields a poor payback in Z-2 when the proportion of zeroes is important, what happens in the opposite case of removal of all zero values in accordance with Z-3? As a reliable 4-traffic model is not feasible due to sample shrinkage, we are left in Column 5 of Table 9 with results for 2 traffics that are fully consistent with Jansson's comment to be quoted below: near linearity. To the best of our knowledge, however, such comments are not based on analyses made with more than 2 kinds of traffic (passenger and freight) as we found none in the literature.

Table 9. Preferred model traffic power values under different aggregation and zero handling rules

	Aggregation	**4 Traffics**		**2 Traffics**			**1 Traffic**
	Z-rule	**Z-1**	**Z-2**	**Z-1**	**Z-2**	**Z-3**	**Z-1**
Traffic T	**Box-Cox transformations and their unconditional *t*-statistics with respect to 0 and 1**						
Intercity trains (GL)	**λ estimate**	**0,377**	**0,445**	**0,83**	**0,704**	**1,008**	**0,779**
	t-stat. w.r.t. 0	[1,39]	[1,86]	[0,83]	**[1,80]**	[0,55]	[0,10]
	t-stat. w.r.t. 1	[-2,31]	[-2,32]	[-9,13]	**[-0.76]**	[0,04]	[-0,03]
Regional trains (TER) & (IdF)	**λ estimate**	**1,114**	**6,194**				
	t-stat. w.r.t. 0	[4,12]	[1,94]				
	t-stat. w.r.t. 1	[0,42]	[1,63]				
Freight trains (F)	**λ estimate**	**3,458**	**3,589**	**0,486**	**-0,479**	**1,260**	
	t-stat. w.r.t. 0	[1,68]	[1,21]	[0,475]	[-0,64]	[1,09]	
	t-stat. w.r.t. 1	[1,20]	[0,87]	[-0,49]	[-0,99]	[0,22]	
Log Lik. value (not rounded)		***-12 866,3***	***-12849,5***	***-12903,6***	***-12 900,9***	***-12 396,4***	***-12 921,4***
Estimated parameter total		***51***	***60***	***35***	***37***	***35***	***29***
Ordinary β_k coefficients		41	41	26	26	26	21
Associated dummy coefficients		0	9	0	2	0	0
Box-Cox powers		10	10	9	9	9	8
Pseudo R^2 (E)adjusted for D.F.		***0,813***	***0,836***	***0,744***	***0,754***	***0,756***	***0,733***
Pseudo R^2 (L)adjusted for D.F.		***0,928***	***0,930***	***0,923***	***0,924***	***0,923***	***0,921***
Number of observations		***967***	***967***	***967***	***967***	***928***	***967***
Column number		**1**	**2**	**3**	**4**	**5**	**6**

Note: although the unconditional *t*-statistics presented here do not suffer from invariance problems, they are computed from first partial derivatives of the maximized Log Likelihood in accordance with Berndt *et al.* (1974) and are therefore less precise than they would be if they had been calculated from second derivatives. Likelihood ratio tests remain strictly precise. The pseudo R^2 measures of goodness-of-fit are calculated according to Equation (10) and Equation (11).

[43] If 1,114 is the true value of the BCT, it is quite normal that the *t*-statistic with respect to 1,000 indicate that 1,114 is not very different from 1,000 in view of the fact that *t*-tests have low power if the values of H_0 and H_1 are close; however, if the true value is 6,194, one would expect much greater certainty that it differs from 1,000 than the indication of 1,63 provided. Clearly, the Log Likelihood surface is quite flat around the BCT value of 6,194 but quite peaked around that of 1,114. Something of the same nature also occurs with the BCT for freight trains. Hence our preference for the result obtained under Z-1.

Choice of a preferred 4-traffic model under Z-1. Form, aggregation and zero value handling rule dimensions therefore all point to the (greyed) U-GBC case of Column E of Table 8 as most performing model for our problem.

In view of the rough equivalence between LL and TL forms, written symbolically [LL ≡ TL] above, it might then be thought that the real contest is not between the TL and U-GBC forms but between the latter and Log-Log forms. Note again that, in the first line of Table 8, the comparison between them indicates a difference of 262,9 Log Likelihood points (for 32 degrees of freedom, on average more than 8,22 points per additional parameter!), a number extraordinarily, not to say infinitely[44], favourable to the U-GBC.

Further work on this preferred U-GBC model involved three different sets of probes. First, the sample for France was split into *Île-de-France* and *Province*: one could reject the hypothesis that Île-de-France results differed significantly from those for the rest of the country, but not the converse view. It was also found, as before with the simpler models of Table 5, that a case could be made for some remaining small regional administrative differences. Finally, the construction of an ***Equivalent Traffic Load*** variable, obtained by weighing train types by segment as specifically recommended by the UIC maintenance manual (UIC, 1989)[45], provided a far less convincing[46] explanation of the maintenance cost than that provided by the preferred U-GBC form.

C. The preferred 4-traffic U-GBC model and the specificity of damages by train category

Estimated BCT and the bowels of the chosen U-GBC model. Detailed results for the 5 competing forms of line 1 of Table 8 can be found in Appendix 3 which is easy to read except for BCT indicators found in Part I of the table, in the two GBC columns, below the *t*-statistic of transformed variables. The presence of a variable-specific λ_k is denoted by LAM and, if the same λ_k is used for many variables, their group number appears, *e.g.* LAM 2. As this is per force complicated, it may be useful in Table 10 to extract the detailed structure and values of the BCT for the preferred case, because there were numerous possible ways to specify the U-GBC form. We note in Table 10 that:

i) **Interactions as product or ratios of variables?** none of the Q*T or T*T interactions is logarithmic except between Q_2 and all traffics, where the BCT is equal to 0,00 [0,004 in fact]. In particular, *ratio interactions* between Q_3 or Q_4 (see Appendix 1) and all traffics *are well identified*, because the relevant BCT, respectively equal to 2,13 and 1,59, strongly differ from the logarithmic case where such identification would not be feasible;

---

[44] Of course, if one degree of freedom were involved, a gain of 8,22 Log Likelihood points would be acceptable at the 99,9999% level. As additional lost degrees of freedom require proportionately lower gains per degree to maintain the same level of certainty, the U-GBC model effectively dominates the TL model with 100% certainty.

[45] The formula for Equivalent traffic load in tons { ^{E}Tr } weighs total gross passenger train tons T_p and total gross freight train tons T_f as follows: { ^{E}Tr } = $\{s_p\,[(1+k \bullet e_p)\,T_p + s_f\,[(k_f + k \bullet e_f)\,T_f]\}$, where the coefficients represent the relative damage of freight axles (k_f = 1,20), engine-specific damage (k = 1,40), the proportion of train tons accounted for by engines (e_p = 0,20 for passenger trains and e_f = 0,10 for freight trains) and the two remaining speed coefficients, defined by segment, vary with the maximum allowed speed on the segment. For passenger trains, s_p increases in unequal steps from a reference level of 1,00 at speeds up to 60 km/h. until 1,50 for speeds greater than 250 km/h.; for freight trains, s_f increases in the same way as s_p up to a maximum of 1,15 because freight train speeds do not exceed 100 km/h.

[46] In a representative test with the LARGE database, the Equivalent traffic variable and its corresponding interaction variables were added to a U-GBC variant based on 4 traffics in order to form an artificial joint model of 71 parameters. Removal of the 5 variables linked to Equivalent traffic reduced the joint Log Likelihood value (of -15240) by 4 points (to -15244, with a difference of 10 degrees of freedom) and removal of the 30 variables linked to standard traffics reduced it by 119 points (to -15321, with a difference of 40 degrees of freedom).

iii) **Own power terms:** in so far as T terms are concerned, it is difficult to be surprised that their estimated BCT are all above the logarithmic value, unless one had TL expectations, because the real issue is rather whether there is here a “power law” comparable to that relating road vehicle axle weights to road damages. We find that the power is lower than 1 (at 0,38) for intercity trains T_1, slightly above 1 (at 1,11) for the regional and suburban trains T_2 and T_3, and *strongly above 3 (at 3,46) for freight traffic* T_4.

In addition to the own-power effect, the total effect of gross train weight depends also on that arising from the interaction terms. Among the 16 terms corresponding to Q*T, the 5 terms that have *t*-statistics above 2,00 all have positive signs ─which implies additional damages arising from the interaction─ and *3 out of those 5 belong to freight trains* (they are doubly framed in Table 9). The interactions arising from T*T interactions are not to be entirely neglected, but are much less significant than the Q*T ones.

Table 10. Values of the 10 Box-Cox transformations estimated in the preferred U-GBC form

C

	C
y	0,31

S

	S
S_1	0,11
S_2	0,11

Q

	Q_1	Q_2	Q_3	Q_4
Q_1	0,11			
Q_2		0,11		
Q_3			0,11	
Q_4				0,11

Q*T

	T_1	T_2	T_3	T_4
Q_1	0,57	0,57	0,57	0,57
Q_2	*0,00*	*0,00*	*0,00*	*0,00*
Q_3	2,13	2,13	2,13	2,13
Q_4	*1,59*	*1,59*	*1,59*	*1,59*

T*T

	T_1	T_2	T_3	T_4
T_1		0,74	0,74	0,74
T_2			0,74	0,74
T_3				0,74
T_4				

T

	T_1	T_2	T_3	T_4
T_1	0,38			
T_2		*1,11*		
T_3			*1,11*	
T_4				3,46

LEGEND	
S_1= Switches	
S_2= Length segment	
Q_1= Rail age	T_1 = GL
Q_2= Sleeper age	T_2 = TER
Q_3= Maxim. speed	T_3 = IdF
Q_4= % concr. sleep.	T_4 = F

Source: Appendix 3, Part II, Column U-GBC, Lines 0-10.

There is therefore strong evidence of a power above 1 for freight trains, but we will see below that their marginal costs, which take all (first order and interaction) effects into account, are low. We will try to make sense of marginal costs that are low while (first-order) damages are strongly rising, in our discussion of marginal costs (calculated with first and second order terms) by train type.

Is there a rail axle load damage power function? At face value, the results of Table 10 argue at least for traffic-specific effects and the presence of a (first-order term) power much greater that 1,00 for freight trains, as did results found in the pared down exploratory model[47]. This differs markedly from some published views, for instance by Jansson (2002):

“*Those studies which have tried to establish a direct link between maintenance cost and axle load have arrived at an almost linear, or slightly progressive, relationship. Although it is not possible to draw any*

[47] In the exploratory model with 5 traffics, we compared specifications where traffic in measured in train-tons, as it is here, with other specifications splitting train-tons between the number of trains and their average weight. Assigning a first and common BCT to passenger train average weights and a second one common to freight and maintenance train average weights yielded a value equal to 1,04 for the former and equal to 7,80 for the latter parameter ─and powers smaller than 1,00 for the 5 BCT associated to number of trains by category.

firm conclusions about how reconstruction and maintenance costs increase with axle load, it is clear that the relationship is far less progressive than the so-called "fourth power law" in the road sector."

who refers directly to the famous "fourth power" axle weight damage law first formulated in the official analysis (HRB, 1962) of the American Association of Highway State Officials (AASHO) Road Test conducted[48] in Illinois between 1958 and 1962. Since then, the prevailing measure of this deterioration is the relative pavement damage of different axle loads. In consequence, the equivalent single axle load rating ("ESAL") or load equivalent factor ("LEF") of an axle or axle group configuration *x* is defined as the number of passes of the standard (reference) axle load required to create the same amount of damage as one pass of candidate axle load *x*:

$$ESAL\,(x) = \left[\frac{Axle\ load\,(x)}{Reference\ axle\ load} \right]^4 \tag{17}$$

The seminal but rather simple LL form AASHO analysis was later redone with the same data and two sophisticated estimation procedures: for fixed Log form models, a one-limit Tobit statistical formulation[49] to account for the censoring of vehicle passes (upwards, at 1 114 000) implied by the limited 2-year duration of the experiment; and, for the authors' strongly preferred Box-Tidwell models, a non linear maximum likelihood procedure implemented in TSP software[50]. In these latter Box-Tidwell tests, the re-estimated optimal load power[51] was closer to 3 for rigid pavements made of Portland cement concrete (Small & Zhang, 1988) and varied between 3,3 (for light loads) and 4,6 (for heavy loads) for flexible pavements (Zhang, 1990).

There is to-day an important literature on appropriate powers for pavements that takes into account varying traffic conditions (*e.g.* traffic mixes ignored in the ASSHO tests, "stop-and-go" traffic[52]), tire width (Lloyd & Addis, 2000), and exacerbating climatic conditions[53]. Climate interacts with heavy traffic[54] in the absence of which significant aging does not occur: to-day, pavement deterioration prediction models cannot separate the load-associated deterioration from the environment deterioration causes. Naturally, national highway construction manuals incorporate explicit values of the axle weight powers. In the Anglo-Saxon world, the AASHO fourth power is

[48] In the main portion of the tests and at a cost then of more than $ 250 million, trucks of 10 distinct types and loads made over 2 years some 1 114 000 passes over 274 sections of road, each constructed to a different specification, The official analysis relates resulting pavement condition to the load, axle configuration and pavement structure but also contains some seasonal weights.

[49] In effect a special case of the RHS of Equation (8).

[50] To the extent we could determine, TSP's maximum likelihood procedure used (VAX version 4.1) automatically calculates asymptotic unconditional values of standard errors of all parameters without due regard to whether some parameters are BCT powers, which implies resulting unconditional *t*-statistics for the two BCT estimated in both papers, *i.e.* values dependent on units of measurement of the transformed pavement and traffic variables. Log likelihood values produced by TSP are straightforward and there is little doubt here that the optimal powers are correct, even in the absence of guarantees of a global maximum, because local maxima are rare, if not unknown, with only two BCT. A formal two-dimensional grid search (on λ_1 and λ_2) would have provided both that certainty, with a plot of likelihood contours, and invariant conditional standard errors or *t*-tests.

[51] These authors also attempted a TL specification but the "wildly varying results for pavement structure and light loads" led them to conclude that this specification was "unreliable" (Small & Zhang, p. 17).

[52] A professor at the Technical University of the German army next to Munich considers that, in stop-go traffic, the power is close to 20 and that it is closer to 3,8 in free-flow steady traffic on a good highway.

[53] As stated by Small (1990) in a summary paper, "In the case of rigid pavements, there is no evidence of significant aging effects, and there are examples of auto-only roads in severe climates that have lasted well beyond 25 years without resurfacing. […] Time and weather exacerbate ─rather than duplicate─ the effects of heavy traffic."

[54] Models where the climate intervenes additively to explain deterioration are not credible, as was admitted on September 26, 1991, by Bruce Hutchinson upon questioning by the Canadian Royal Commission on National Passenger Transportation during the discussion of a model by Jung *et al.* (1975) that treated climatic factors additively (in Hutchinson, 1991). The treatment of weather and traffic effects as additively separable is now outdated, hopefully.

still the practical reference[55]. In France, the power 5 is used for flexible pavements and 12 for rigid and concrete pavements (LCPC & SETRA, 1994). All such values[56] are now derived from laboratory tests rather than the great outdoors.

There is also a derived literature on joint optimization of pavement construction and maintenance, called life-cycle cost analysis, where pavement thickness and maintenance can be jointly determined optimally (*e.g.* Small *et al*., 1989). Optimization is critical because, as pointed out by Heggie (1991a)[57], "while the structural damaging effect of traffic rises with the fourth power of axle loadings, pavement strength rises with a typically ninth power of thickness and cost (a power of about 5 to 13, depending on materials)". The demonstrated existence of power laws for rail tracks could therefore modify current belief and practice apparently reflected in Jansson's summary view.

Conditions needed to detect rail axle load power functions. We therefore have indications that the establishment of power functions for the rail mode may just require more than 2 traffics but would prefer data on individual train weights and suspension characteristics to pursue the issue. It is also the case that, as demonstrated in Appendix 1, it is critical to determine damages for given quality of track, as done in the AASHO tests for different pavement qualities but yet to be effected convincingly in rail studies.

The demonstration of the presence of power laws should be facilitated by the flexibility of BCT: simple power parameters are not good substitutes because terms such as y^{γ} can lead to the degenerate solution $\gamma = 0$ as one approaches power values that matter but differ from the logarithm: a BCT solves this problem and avoids being obliged to use for the dependent variable y an priori form such as the logarithm, as was done in the re-estimations of the AASHO tests.

The analysis of the behaviour of the preferred model under aggregation therefore provides strong evidence of train-specific weight damage effects in accordance with (H_0-2) and even of the presence of rail "power laws" analogous to (17) requiring further confirmation, preferably with data on individual trains and their composition.

D. The preferred 4-traffic U-GBC model and cost recovery: other non linear cases of Table 8

Our preferred model has been chosen primarily on the basis of an analysis of Log-Likelihood values analysed by line and for all lines of Table 8 and of a detailed comparison among all U-GBC model values from Column E. Will this preference be maintained as we consider cross-cases from the rest of Table 8 columns except for the first (neglecting linearity for obvious reasons), and shift the emphasis from statistical to economic considerations, notably with respect to maintained hypothesis (H_0-3) on cost recovery from hypothetical marginal cost charges?

Cost recovery and form when zero values are treated. The influence on cost recovery from hypothetical marginal cost pricing regimes of form and Z-rules is best summarized in Figures 1 and 2 where the X axis shows the number of traffic categories used and the three lines indicate hypothetical cost recovery from marginal cost charges determined by each model form.

In Figure 1, the adjustments are drawn from Part I results of Table 8 (Log-Log, Trans-Log and U-GBC): the recovery ratio drops dramatically with the number of traffics if pricing is derived from Log-log and Trans-log forms. *This does not happen with the Box-Cox transform* which is

[55] Wyn Lloyd of the United Kingdom Highways Agency thinks that the fourth power law is viewed in the U.K. as a « reasonable compromise » between estimates ranging from the power 3 to the power 8 (Lloyd & Addis, 2000).
[56] For bridges, the power 3 of Gross Vehicle Weight, not of axle weight, is often deemed appropriate (Fekpe, 1996).
[57] In his Annex 4, not found in the shortened version (Heggie, 1991b).

insensitive to the zero values of variables *provided that the powers of the transform be positive*, which is always the case for our sample. This may be due to segments where some traffic is null as occurs for 39 segments with the 2-traffic specification, much more with the 4-traffic specification (for instance on many segments there is no Île-de-France traffic) and never with the 1-traffic specification. Our tests performed on both the SNCF sample and on simulations drawn from random numerical samples have shown that cost coverage for LL and TL forms is very sensitive to the small arbitrary 0 replacement value.

Figure 1. Cost recovery under marginal cost pricing with models from Table 8, Part I (Z-1 rule)

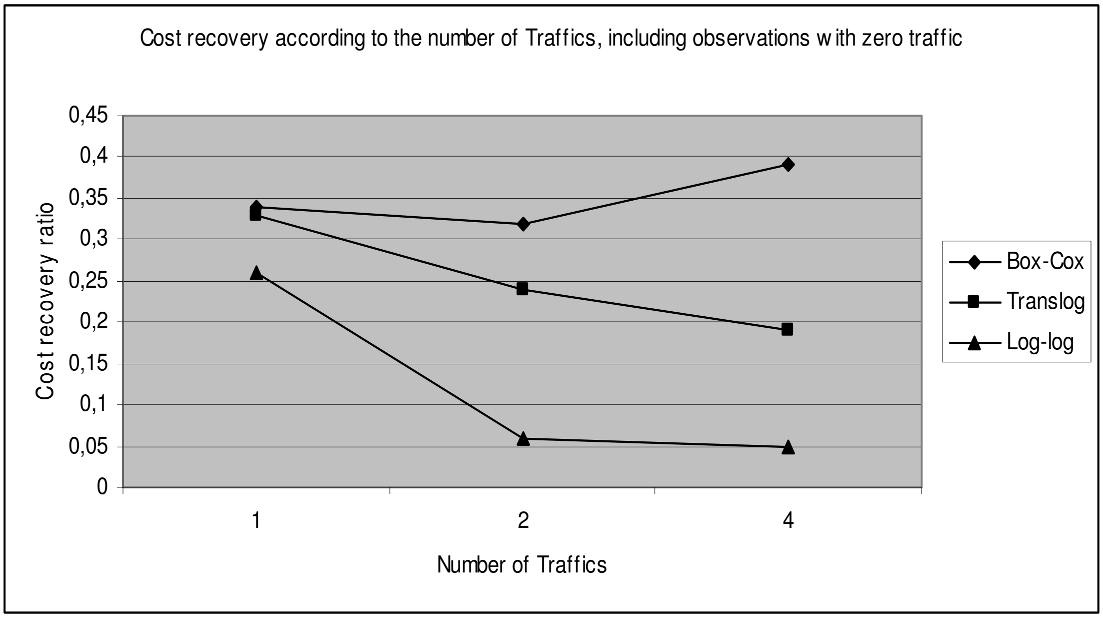


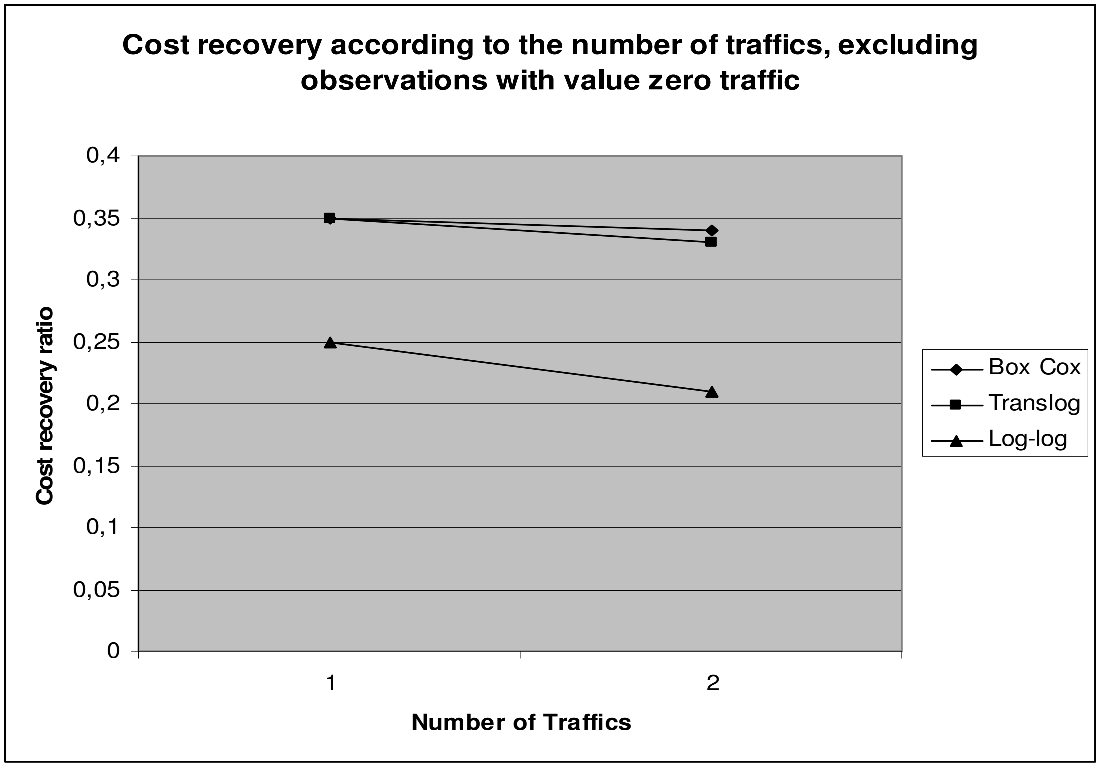
Figure 2 confirms this hypothesis by using Part III results (Log-Log, Trans-Log and R-GBC) derived from strictly positive observations: the set of observations is slightly reduced from 967 to 928 to exclude the 39 observations for which there are some zero value traffic. It is clear in this case that using one or two traffics makes much less difference to cost recovery rates.

Cost recovery and form in the absence of zero values. The Figure 2 results also favour Box-Cox specifications in other ways. We have just noted that BCT are more robust to the range of traffic values than Log-log or Trans-Log specifications and to the exact replacement value used for the output values equal to zero. Now we note that there may be a third reason for maintaining our preference for the Box-Cox models that generally support higher numbers of traffics (less aggregation): as we look again in Figure 2 (where zeroes do not interfere), we find that cost recovery is higher for Box-Cox adjustments irrespective of the number of traffics. This result was also found by Link (2009) for Austrian railways with a single-output model comparison between a LL form and its BCT generalization.

Preference for U-GBC (or eventually for R-GBC) models is thereby reinforced: they favour disaggregation by user class in accordance with (H_0-2), are more stable than other models in the presence of zeroes and, concerning (H_0-3), yield higher cost recovery levels than other forms.

Aggregation and cost recovery. Is this last result to be expected? Considering also the unproblematic Box-Cox adjustments in Figure 1, it appears that the cost coverage is higher for the 4 traffics variant than for the 1 or 2 traffic ones, thereby implying again that the Box-Cox models imply higher recovery rates. This double result makes sense, as we presently argue.

Figure 2. Cost recovery under marginal cost pricing with models from Table 8, Part III (Z-3 rule)



Should the cost be explained just by traffics (or should the other variables be uncorrelated with traffics) and should the relation be linear, the adjustments would be written, successively for the "one-traffic" and the "two-traffic" models as:

$$C = b_1 + a_1*(T_1 + T_2) + e_1 \tag{17}$$

$$C = c_2 + a_2*T_1 + b_2*T_2 + e_2 \tag{18}$$

where T_1 and T_2 denote traffic types. It is clear that the multiple R^2 correlation coefficients necessarily verify $r_2>r_1$ because of the restriction $a_2=b_2$ imposed in the first model, as compared with the second. It is also clear that, as both expressions are positive, we have by the properties of Least Squares $\sum(a_2*T_1 + b_2*T_2) > \sum a_1*(T_1+T_2)$, which implies that the recovery rate with 2 traffics is higher than with 1 traffic and, similarly, is also higher with 4 traffics than with 2 traffics.

The differences between this simple case and those of Table 8 are that the latter relations are not linear and that their other variables are somewhat correlated with their traffics. Nevertheless, the result holds as long as the relation is not "too" non linear and the other variables are not "too" correlated with the traffics: it gives a hint that recovery rates are presumably larger with 4 than with 2, and with 2 than with 1, traffic.

This ordering is similar to that suggested by a comparison of Box-Cox model goodness of fit statistics across levels of traffic aggregation. First, the Pseudo-R^2-L values (greyed in Table 9) duly increase with the number of traffics considered under the Z-1 rule; second, even if the marginal gains somewhat decrease with increased numbers of traffics, the absolute gains in Log likelihood are still important: 17,8 points from Column 6 to Column 3 [with 3 degrees of freedom lost] and 37,3 points from Column 3 to Column 1 [with 16 degrees of freedom lost].

Full models and the three maintained hypotheses from exploratory work. All results of Table 8 considered, our preferred model choice confirms maintained hypotheses concerning form (H_0-1), damages varying by user class (H_0-2) and differentiated cost recovery under a marginal maintenance cost recovery pricing *régime* (H_0-3). Very detailed full models do not modify the overall bearing of the results obtained with exploratory pared down models.

5.3. Quality-Traffic interactions and model performance: products, but not of logarithms

But the extreme generality of the U-GBC form makes it difficult to intuit the reasons for its performance relative to that of other forms, and notably relative to the TL form, long-established as the workhorse of total and maintenance cost studies in rail transport.

Simply put, if the TL, which contains interaction terms, is no better than the LL, which contains none, the source of the advantage of the two GBC variants over all fixed form specifications must be in the flexibility of the BCT, *i.e.* in modelling the interactions better than in the TL, a diagnostic confirmed by the further superiority of the U-GBC over the R-GBC.

To understand this better, we carefully examine the role of technical variables $\boldsymbol{K} \equiv \boldsymbol{(S, Q)}$ considered by themselves (for first order effects) and considered also in their interaction roles. We isolate the role of Q*T interactions and demonstrate that their presence does not guarantee their having a role unless they are introduced with the correct form. This can be seen if we study Q*T interactions in the 4-traffic models under Z-1 or if we take some of these models and systematically remove interactions from them in step-wise fashion. In both cases, Q*T interactions matter only if they are inserted in the correct form and not as products of logarithms.

A. Studying variables, including interactions, in the 4-traffic models under Z-1

State variables. The State variables are very significant, as can be seen in Table 11 where segment length may reveal the presence of unexploited economies of scale because the explained cost is defined *per kilometre*. Interesting enough, but our concern is with the interacting variables.

Own terms : negligible Q*Q interactions, moderate T*T interactions. Quality variables are a more complicated problem than State variables because of their potential interactions among themselves (Q*Q) and with Traffic variables (Q*T). We mentioned above that Q*Q interactions had been removed from Box-Cox specifications because they were never significant in these forms: the same result is found in the TL form where they were retained in principle, as may be readily verified in Appendix 3 where it is also possible to see that T*T interactions are much more significant, notably between T_3 and T_4 than the Q*Q interactions. In the Box-Cox forms, the most important interactions by far are the Q*T interactions.

Table 11. Results for selected State variables (reference model series)

Cost elasticities with respect to Switches and Segment Length (I.4 Series A: [4 Traffics] in Table 8)					
Form	**LIN**	**LL**	**TL**	**R-GBC**	**U-GBC**
Number of switches					
Elasticity evaluated at sample means according to Eq. (13)	0,45	0,28	0,26	0,28	0,24
Conditional *t*-statistic of β_k coefficient	(18,12)	(10,42)	(8,92)	(12,63)	(10,51)
Segment length					
Elasticity evaluated at sample means according to Eq. (13)	-0,29	-0,30	-0,29	-0,28	-0,26
Conditional *t*-statistic of β_k coefficient	(-9,98)	(-11,93)	(-11,00)	(-11,69)	(-12,57)
Source: Appendix 3, Part I, Line 1 (variable *apdv*) and Line 4 (variable *long*).					

Q*T interaction in Appendix 3. However, testing the role of interactions in the U-GBC form is not straightforward because this format allows many refinements excluded by the R-GBC restrictions. For instance, redefining the Q*T interactions, characterized by 4 *horizontally* constrained BCT in Table 10, by 4 *vertical* restrictions gave much less good results: Q*T interactions are therefore better defined by making them vary by quality independently from train categories rather than by train type independently from the nature of qualities. As flexible horizontal interaction constraints are much better than flexible vertical ones, the imposition of TL logarithmic symmetry in Appendix 3 yields a predictably mediocre outcome: wrong in its symmetry assumption and wrong in the value of the imposed uniqueness of the implied power term.

Fair enough, but we did not explore the most general specifications that would allow 16 distinct interactions by (λ_q , λ_p) pair in (U-GBC) as this would have increased by 12 [= 16 - 4] the number of estimated BCT (symmetry of the matrix is not meaningful here) and forced time consuming tests to guarantee global maxima of the Log Likelihood function: we had already found 2 or 3 examples of local maxima with 9 to 11 BCT parameters. It is a matter of time before algorithms compute meta-statistics on the probability of having obtained a global maximum, but our L-1.4 algorithm required that the raw data for such robustness statistics be derived by tedious manual modifications of the starting values of the vector of BCT parameters, a non trivial problem with say 21 to 23 BCT.

B. The role of quality traffic Q*T interactions identified by step-wise addition of groups

Another reasonable way to understand the importance of the interactions involving the quality variables is to add them in progressively, as done in the following Table 12 where, in TL and R-GBC forms: one starts, in Case C, with only Traffic and State variables[58] and successively adds, in Case B, the four Quality variables (and the weaker T*T interactions) before finally adding, in case A (drawn from Table 8), all interactions among Quality and Traffic variables.

Table 12. Effect of Quality, Traffic-Traffic and Quality-Traffic variables

		TL[(1)]			R-GBC		
		Log Likelihood	Parameters estimated	Elasticity of cost	Log Likelihood	Parameters estimated	Elasticity of cost
4 traffics (Rule Z-1, 967 observations)							
Case C	T, S	n.a.	--	--	-12978	16	0,37
Case B	T, S, Q, T*T	-13104	39	0,18	n.c.([3])	--	
Case A[(2)]	T, S, Q, T*T, Q*T	***-13093***	***55***	***0,19***	***-12906***	***42***	***0,38***
2 traffics (Rule Z-3, 928 observations)							
Case C	T, S	n.a.	--	--	-12475	14	0,23
Case B	T, S, Q, T*T	-12598	30	0,32	-12433	18	0,37
Case A[(2)]	T, S, Q, T*T, Q*T	***-12593***	***38***	***0,32***	***-12421***	***29***	***0,34***
1 traffic (Rule Z-3, 928 observations)							
Case C	T, S	n.a.	--	--	-12468	13	0,24
Case B	T, S, Q, T*T	-12595	27	0,34	-12433	17	0,34
Case A[(2)]	T, S, Q, T*T, Q*T	***-12592***	***31***	***0,34***	***-12421***	***22***	***0,36***

Parameters estimated include both regressors and Box-Cox transformations. For Series A, see Appendix 3, Part III, Line 5.
Elasticity : mean elasticity of the cost with respect to traffic, according to Equation (15), weighted by traffic shares. Note that higher elasticities generally imply, by Equation (16), higher marginal costs.

[(1)] In all A cases, the TL form includes Q*Q interactions but the GBC does not, as pre-specified in Table 4.
[(2)] Case A results in bold italics contain new information not presented in Table 8 (Part I and Part III) but found in Appendix 3.
[(3)] Does not converge.

These additions present interesting features for the understanding of the role of Q*T interactions:

- **Any Q*T interaction will not do.** Comparing B and C cases, Quality-Traffic interaction variables Q*T improve the quality of the adjustments only slightly in the Trans-Log case but considerably in the Box-Cox cases, which suggests again that the *proper form of interaction matters* and that a

[58] Case C of the R-GBC specification is the closest we have to the tests made by Andersson (2009a, 2009b), Link (2009) and Marti *et al.* (2009) on single-BCT models containing qualities Q but no interaction terms. In three of the four cases considered by these authors, the Box-Cox model was preferred to the LL model.

simple product of logarithms will not do. Clearly, many interactions specified *a priori* in the Trans-Log are unrealistic in a physical model of track deterioration, such as those among qualities Q*Q that never seem to matter here, irrespective of form, and those that may be relevant but not in a logarithmic shape. The GBC, preferably in U-GBC garb, is the only adequate specification: imposing a TL form even worsens the adjustment, as compared to a LL.

- **Q*T interactions included in short-run cost function.** Reading the table upwards from the bottom A lines, which show the most complete adjustments in terms of groups of variables, the cost elasticity with respect to traffic decreases when the quality-traffic variables are dropped. This point would deserve more thorough investigation along the following lines: the A and B adjustments with Traffics, State and Quality-Traffic variables together can be seen as representing the short run cost function, while the C adjustments neglecting qualities Q and Q*T interactions can be seen as closer to the medium run cost function. The fact that short run elasticity (and whence the implied marginal cost) in A and B cases is higher than the medium run elasticity (and implied marginal cost) in C cases supports the Rivier & Putallaz (2005) diagnostic that SNCF track renewal policy is under-optimized.

- **Q*T and other statistically significant interaction terms and cost recovery**. Fourth, taking interactions into account in the Trans-Log cases makes no difference to the elasticities. This makes sense: if the Trans-Log is insufficiently parsimonious, in the sense that it includes too many second degree interaction terms (of fixed and incorrect form) and that removing those makes little difference to the quality of adjustment, then we do not expect terms with coefficients barely different from zero to have much impact on derived calculations of elasticities (or of marginal costs). By contrast, in the Box-Cox cases where (flexible) interactions have a more statistically significant role, the elasticities do change.

We conclude studying the role of interactions, and in particular of Q*T interactions, across model forms, that *having the wrong form of interaction is no better than having none* and that the right interaction form cannot be determined *a priori*. The point is of interest because the effective choice in many papers is between fixed parsimonious LL forms and fixed profligate TL forms (e.g. Link *et al.*, 2008): we are effectively arguing that this is the wrong choice due to the lack of flexibility of the TL fixed form in handling interactions terms, rather than due to its generous parametrization.

5.4. Solving a European cost recovery puzzle

Solving a European cost recovery puzzle. These analyses throw some light on the oddness of cost recovery results for France (around 35%, as indicated in Figures 1 and 2) in contrast with average European results (between 15 and 20%, as reported in Wheat & Smith (2008)).

We have seen that Box-Cox forms tend to yield higher recovery rates than those found with other mathematical forms, as does the break-down of traffic variables into four traffic services (most European studies use only one). It is also the case that multiple quality indicators (most European studies have fewer, if any) are relevant because cost recovery falls when they are dropped from the regression. These combined effects explain the “French exception” as they imply a range where the lowest value reaches the higher values of the acknowledged European range, as shown in Table 13.

Table 13. Range of recovery rates in our non linear results

	LL		TL	R-GBC
Case A. 4 traffics with Q*T variables	0,05*		0,19	0,39
Case A. 1 traffic with Q*T variables	0,26*		0,33	0,34
Case C. 1 traffic without Q*T variables	0,18**		0,24	n. a.
Source	*Figure 1	**Other	Table 12	Table 12

There is less to learn immediately from the system-wide elasticities and corresponding marginal costs listed in Table 12: they will be analyzed further below to understand their components across traffic categories and their values elsewhere than at the means.

6. Detailed elasticities and marginal costs derived from the preferred model

Up to this point, only mean *system-wide* values of elasticities, marginal costs and implied cost recovery ratios have been used in our comparison of non linear cases in Table 8 (*e.g.* Figure 1, Figure 2 and Table 13)[59]. But mean values *by user class* are also of interest, as are evaluations at *loci other than sample point* values: in the case of monotonically changing relationships, a mean evaluation, for the system or a user class, is fine as it goes but becomes misleading if the underlying relationship is U-shaped. We therefore extend our analysis of the impact of traffic on cost beyond mean system-side measures: first towards mean values by user class and second towards values by user class, considered one at the time and jointly, obtained by simulating*, ceteris paribus*, traffic loci situated in representative domains[60] of the four traffics, other variables held at sample means.

6.1. Average value by service, at traffic sample points

Table 14 presents the averaged cost elasticities and marginal costs with respect to traffic calculated by traffic type and for the system, respectively.

Table 14. Averaged marginal cost and cost elasticity with respect to sample traffic service points

Averaged point marginal cost (in Euro per 100 ton-km per year) and elasticity with respect to traffic					
	T_1 Intercity	**T_2 Regional**	**T_3 Ile-de-France**	**T_4 Freight**	***System***
Marginal cost	0,172	0,458	0,174	0,069	***0,139***[(1)]
Elasticity	0,118	0,122	0,033	0,114	***0,387***[(2)]
[(1)] Average of "averaged" values obtained according to Equation (16). [(2)] Sum of values for the four individual traffic categories, each obtained according to Equation (15).					

Unexpected marginal freight costs. A noticeable feature of this set of results is that the marginal cost of freight is lower than the marginal cost of all passenger traffics. This point, paradoxical to the extent that engineering approaches conclude that freight trains are more damageable than passenger trains, might be reasonable if maintenance costs depended not only on the damages but also on the cost of maintenance works, on the required level of quality, and on available funds. Successively:

- **Maintenance work conditions.** The first reason is the "track possession" time allowed for repairs. Works are more expensive when track possessions are short, and this is the case in France in day periods for passenger trains whose timetable cannot easily be adapted to needs of maintenance while freight train time–tables are easier to change: this point is specially relevant in France where very short track possessions are traditionally accepted, a policy that is currently changing. Of course, track possession are specially short on segments with a large numbers of trains (which is the case of tracks with regional trains) and induce more expensive maintenance, except when the maintenance is achieved during the night, which is rarely the case in France.

- **Maintaining quality.** The maintenance cost depends not only on the damages caused by the trains but also on the quality level objective; in that respect it is clear that segments with a large proportion of freight trains do not require a high level of quality while segments with a large proportion of passenger trains require a high quality level. A model of social optimization of maintenance shows that, when the traffic is composed of trains of type A demanding no quality but damaging the track (the case of freight trains) and trains of type B demanding a high quality

[59] An exception is Table 11 where elasticities are calculated at sample means and are not mean elasticities.
[60] This amounts to a manual modification of one or more traffic variable X_k in (13).

but giving little damage to the track (the case of passenger trains), type B trains can have a marginal cost higher than type A trains, as shown in Appendix 2.

▪ **Maintenance postponed.** This quality effect is enhanced when there is a shortage of funds, a fact which leads to favour, in terms of prevention and renewals, the most circulated tracks and increases the marginal costs on other tracks because they are then more subject to curative maintenance, as opposed to cheaper preventive maintenance. But the resulting increases in marginal freight costs appear here to be dominated by the reduced need for maintenance on low quality lines, a factor working in the opposite direction. The net result found in Table 14 implies that the cost minimization assumption is distorted by short run considerations that make the maintenance expenditures more dependent on the traffic "standing" of lines than expected from optimal planning considerations. Clearly, the different quality factors in our models are then insufficient to account for "standing", thereby causing passenger costs to depend more on traffic than they should. This prevalence of standing amounts to a missing variable in the models.

If the problem is as just described, low marginal costs for freight trains are perfectly consistent with higher power damage functions for these trains.

6.2. Beyond mean values: U-shaped or not?

Remember that variables other than Traffic include State and Quality variables at their sample means: this will be left unchanged in the first construct and changed forthwith in the second.

Marginal cost by traffic type. The first case is considered in Table 15 where, interestingly, the marginal cost for long distance passenger trains (Intercity) indicated in the greyed column exhibits a textbook-like U shape with a minimum roughly at mid-point of the traffic level, *i.e.* not too far from the actual mean level of 1 116 095 indicated in Table 3. However, as the own coefficient of the Intercity traffic variable itself is positive (see Appendix 3), this net turning effect seems entirely due to the (additional) interaction terms, somewhat worrying if one expects own effects to dominate.

On this point, Andersson (2006, Table 6) has estimated a 1-trafic maintenance cost model with a strict quadratic form term very significant in the logarithms of his traffic variable (gross tons). By contrast, in a similar model, Marti *et al.* (2009) exclude from their reference model the added quadratic traffic term found to be of marginal statistical significance. The presence of U-shaped effects is therefore far from a standard feature of current models.

Table 15. Average marginal cost and cost elasticity with respect to simulated traffic service points

Simulated traffic				Marginal cost (€/ 100 t-km/year)				Elasticity			
Intercity	Regional	IdF	Freight	Intercity	Regional	IdF	Freight	Intercity	Regional	IdF	Freight
41 921	16 384	11 695	30 000	0,50	1,03	0,49	0,46	0,04	0,03	0,01	0,02
419 207	163 844	116 950	300 000	0,15	0,35	0,20	0,13	0,08	0,07	0,03	0,05
2 096 035	819 218	584 748	1 500 000	0,07	0,43	0,11	0,07	0,11	0,26	0,05	0,08
4 192 070	1 638 435	1 169 495	3 000 000	0,07	0,52	0,05	0,06	0,15	0,42	0,03	0,09
8 384 139	3 276 870	2 338 991	6 000 000	0,14	0,61	-0,09	0,03	0,34	0,56	-0,06	0,05
16 768 278	6 553 741	4 677 981	12 000 000	0,46	0,36	-0,44	-0,27	1,17	0,35	-0,31	-0,49

Simulated total cost elasticity and marginal cost and by track quality. Results in Figure 3 for the second case pertain to three simulations of the effects of joint traffic variations based on different values of State and Quality indicators: average quality, defined by the average values of S and Q, and low or high quality, defined by values of S and Q some 30% lower or higher than the mean ones. Again, the joint effect differs from the effects of traffic considered one at the time. Elasticity results shown in Figure 4 are constructed in the same way with different values of State and Quality indicators. Together, these two figures complement the system-wide results found in Table 14.

Figure 3. Marginal traffic cost as a function infrastructure quality

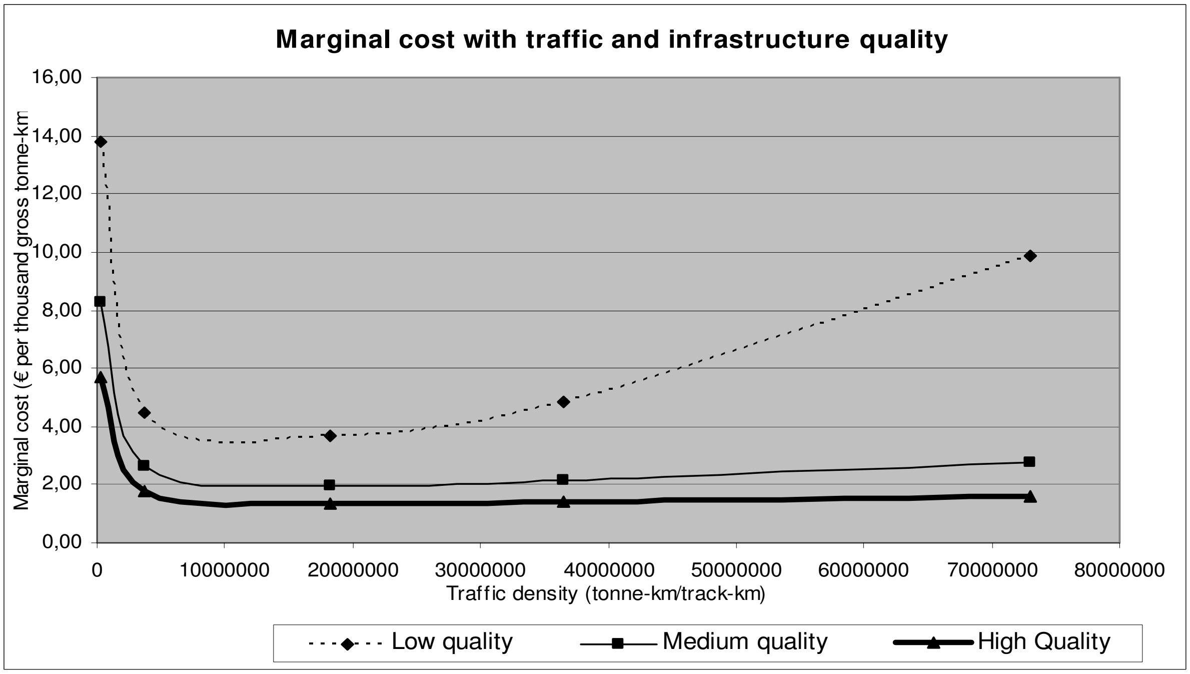
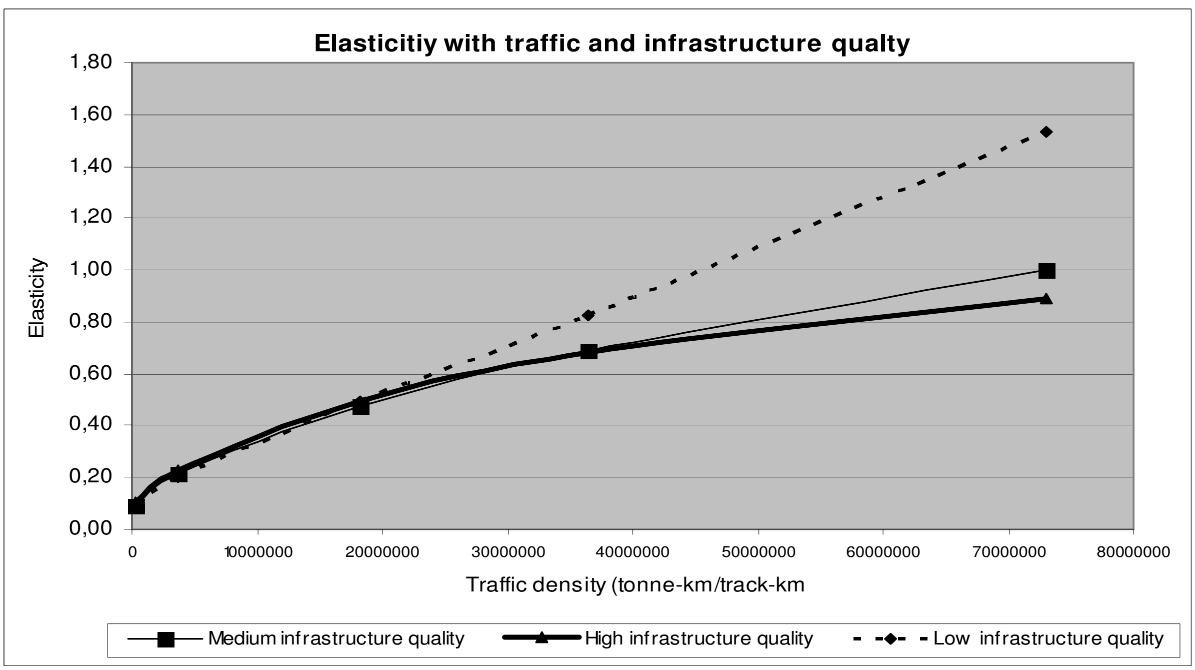


Figure 4. Cost elasticity with respect to jointly changing traffic as a function of infrastructure quality



U shapes an issue. U shapes are therefore quite delicate a topic. In models containing only monotonic terms, the presence of a real, but hidden, quadratic effect can imply very low coefficients of the traffic variables because fitting a straight line, or of a simple monotonic curve, through a U can easily lead to a misleading slope coefficient not too different from zero, and even of the "wrong" sign: it is therefore important to look for quadratic effects by testing explicitly for the presence of turns as part of the core search. But such tests have their own complexities because an *asymmetric* U is obtained if one generalizes the specification ($y = f(X_k, X_k^2)$) by using a BCT for the squared term[61], allowing the estimated power to differs from 2. It is sometimes possible to reject a symmetric form ($\lambda = 2$), when tested against its removal, and to accept an asymmetric form ($\lambda \neq 2$), essentially because a symmetric quadratic is a rigid and limited special case of turning functions.

Overall, it seems that the issue of the presence of U-shaped effects in the analysis of maintenance cost should be the object of a distinct analysis, preferably with data on individual trains. In this study, the Linear and Log-Log forms exclude it from the start and, in the Trans-Log form, the predominance of first order terms is considerable; overall, we found the presence of U-shaped effects demonstrably rare.

7. Conclusion: letting the data, not the analyst, decide the form

Unfortunately, we have no models of maintenance costs to test our GBC results against, and no analyst's "theory" is yet strong enough to predict parameter and form values. In this applied case, the Unrestricted Generalized Box-Cox form gives the best results, provides strong evidence of the existence of power function by train type and, consequently, of user-specific marginal costs.

But, if our primary concern is with the reasonableness of an *ad hoc* specification, it is also with whether the results (*e.g.* on marginal costs) make sense. On this point, we were satisfied with the economic values implied by the "optimal" results and draw some methodological conclusions:

Trans-Log and Log-Log specifications are not convenient if some traffic variables are sometimes null: they do not represent well the behaviour of cost functions drawn from engineering knowledge. From this point of view, the Box-Cox form is innocuous, as is the linear one.

Technical (both Quality and State) variables are highly significant. It is important to use them in the adjustments in order to properly estimate the short run cost function: without them, the function would be more akin to a long run cost function, assuming an optimal capital level.

It appears that interactions among Quality variables and Traffics are significant only in the Generalized Box-Cox cases and especially in the most flexible unrestricted U-GBC case. This point is a bit worrying as these interactions should incorporate the substance of the track, which is acknowledged by the engineers as an important factor of maintenance costs: when the age of the track approaches a level expressed in cumulated ton-km, damages and maintenance costs increase sharply. Unfortunately, the present estimates fail to reproduce this point in the Trans-Log case. Further research is needed on this point, and generally on links between maintenance and renewal.

We conclude that the Unrestricted Generalized Box-Cox form should be decisive in two sets of circumstances between which most real world situations are found: at one extreme, when one has a clear engineering (or other) model of maintenance costs and one is interested in testing the compatibility of data with it; at the other extreme, when one has no clear maintenance cost model form and the primary concern of the analysis is the reasonableness of an *ad hoc* specification in terms of fitting the data and yielding convincing results. In both cases, of course, we allow the data to decide on the proper form in the hope that the associated results will be convincing in the sense that they do not imply too strong a conflict with *a priori* expectations however derived.

[61] For a discussion of the conditions for a maximum or a minimum, the reader may consult Gaudry *et al.* (2000) where the formulation is more general still and allows for two BCT on the two terms found in ($y = f(X_k, X_k^2)$), a generalization that gives great flexibility to the shape of the turning curve.

8. References

Abrantes, P., Wheat P. E., Iwnicki, S., Nash, C. and A. Smith (2007). CATRIN (Cost Allocation of TRansport INfrastructure cost), ***Review of rail track cost allocation studies for Deliverable 1 of CATRIN***. Funded by Sixth Framework Programme. VTI, Stockholm.

Aigner, D. J. (1971). ***Basic Econometrics.*** Prentice-Hall, Englewood Cliffs, N. J.

Amemiya, T. and J. L. Powell (1980). ***A comparison of the Box-Cox maximum likelihood estimator and the nonlinear two stage least squares estimator.*** Technical Report No. 322, Stanford University.

Andersson, M. (2006). ***Marginal railway infrastructure cost estimates in the presence of unobserved effects***. Manuscript, VTI, Stockholm, 11th September 2006.

Andersson, M. (2009a). CATRIN (Cost Allocation of TRansport INfrastructure cost), Deliverable 8, Rail Cost Allocation for Europe – Annex 1A– ***Marginal Costs of Railway Infrastructure Wear and Tear for Freight and Passenger Trains in Sweden***. Funded by Sixth Framework Programme. VTI, Stockholm.

Andersson, M. (2009b). ***Marginal cost of railway infrastructure wear and tear for freight and passenger trains in Sweden.*** Department of transport economics, VTI, Borlänge, 29 pages, 31 March. Presented at the Fourth International Transport Economics Conference, Minneapolis, June.

Baumgartner, J. P. (1978). The Production Function of a Railway. ***International Journal of Transport Economics*** 5, 2, 145-159.

Berndt, E. R., Hall, B. H., Hall, R. E. and J. A. Hausman (1974). Estimation and Inference in Nonlinear Models. ***Annals of Economic and Social Measurement***, 3, 653-665.

Berndt, E. R. and M. S. Khaled (1979). Parametric Productivity Measurement and Choice among Flexible Functional Forms. ***Journal of Political Economy*** 87, 6, 1220-1245, December.

Bereskin, C. G. (2000). Estimating Maintenance-of-Way Costs for U.S. Railroads After Deregulation, ***Transportation Research Record*** 1707, 13-21.

Bickel, P. J. and K. A. Doksum (1981). An analysis of transformations revisited. ***Journal of the American Statistical Association*** 76, 374, 296-311.

Bierlaire, M. (2003). BIOGEME: a free package for the estimation of discrete choice models. ***Proceedings of the 3rd Swiss Transportation Research Conference***, Ascona, Switzerland.

Bierlaire, M. (2008). ***An introduction to BIOGEME Version 1.6.*** www.biogeme.epfl.ch.

Blackorby, C., Primont, D. and R. R. Russell (1977). On Testing Separability Restrictions with Flexible Functional Forms. ***Journal of Econometrics*** 5, 195-209.

Borger, B. de (1991). Hedonic versus homogeneous output specifications of railroad technology: Belgian railroads 1950-1986, ***Transportation Research 25A***, June, 227-238.

Borger, B. de (1992). Estimating a Multiple-Output Generalized Box-Cox Cost Function: Cost Structure and Productivity Growth in Belgian Railroad Operations, 1950-86. ***European Economic Review*** 36, pp. 1379-1398.

Borts, G. H. (1960). The estimation of rail cost functions. ***Econometrica*** 28, 1, January, 108-131.

Box, G. E. P. and D. R. Cox (1964). An Analysis of Transformations. ***Journal of the Royal Statistical Society***, Series B, 26, 211-243.

Box, G. E. P. and P. W. Tidwell (1962). Transformation of the Independent Variables. ***Technometrics*** 4, 531-550.

Brown, R. S., Caves, D. W., and L. R. Christensen (1979). Modeling the Structure and Cost of Production for Multiproduct Firms. ***Southern Economic Journal*** 46, 256-273, July.

Burgess, D. F. (1974). A Cost Minimization Approach to Import Demand Equations. ***Review of Economics and Statistics*** 56, 225-234.

Caves, D. W., Christensen, L. R. and J. A. Swanson (1980b). Productivity in U.S. Railroads, 1951-1974. ***Bell Journal of Economics*** 11, 1, 168-181.

Caves, D. W., Christensen, L. R. and M. W. Tretheway (1980a). Flexible Cost Functions for Multiproduct Firms. ***Review of Economics and Statistics***, 477-481 (Notes Section).

Caves, D. W., Christensen, L. R., Tretheway, M. W. and J. Windle (1985). Network effects and the measurement of returns to scale and density for U.S. railroads. Ch. 4 in A. F. Daughety (ed.), ***Analytical Studies in Transport Economics***, 97-120, Cambridge University Press, Cambridge.

Christensen, L. R., Jorgenson, D. W. and L. J. Lau (1971). Conjugate Duality and the Transcendental Logarithmic Function. ***Econometrica*** 39, 4, 255-256, July.

Dagenais, M., Gaudry, M. and T. C. Liem (1987). Urban Travel Demand: the Impact of Box-Cox Transformations with Nonspherical Residual Errors. ***Transportation Research B*** 21, 6, 443-477.

Davidson, R. and J. G. MacKinnon (1985). Testing Linear and Loglinear Regressions against Box-Cox alternatives, ***Canadian Journal of Economics/Revue canadienne d'Économique*** 18, 3, 499-517.

Denny, M. and M. Fuss (1977). The Use of Approximation Analysis to Test for Separability and the Existence of Consistent Aggregates. ***America Economic Review*** 67, 3, 404-418, June.

Denny, M. and C. Pinto (1978). An Aggregate Model with Multi-Product Technologies. Ch V.1, in Fuss, M. and D. McFadden (eds), ***Production Economics: A Dual Approach to Theory and Applications***, Vol. 2, 249-267, North Holland, 1978.

Diewert, W. E. (1971). An Application of the Shephard Duality Theorem: A Generalized Leontief Cost Function. ***Journal of Political Economy*** 79, 481-507.

Diewert, W. E. (1973). ***Separability and a Generalization of the Cobb-Douglas Cost, Production and Indirect Utility Functions.*** Research Branch, Department of Manpower and Immigration, Ottawa, January.

Diewert, W. E. (1974). Functional Forms for Revenue and Factor Requirements Functions. ***International Economic Review*** 15, 199-130.

Draper, N. R. and D. R. Cox (1969). On distributions and their transformations to normality. ***Journal of the Royal Statistical Society*** B31, 472-476.

Fekpe, E. (1996). Bridge Cost Allocation Issues. ***Canadian Transportation Research Forum 1996 Proceedings***, Volume 2, 883-892, May.

Fuss, M., McFadden, D. and Y, Mundlak (1978). A Survey of Functional Forms in the Economic Analysis of Production. In M. Fuss and D. McFadden, eds., ***Production Economics: a Dual Approach to Theory and Applications***, Ch. II.1, 219-268. New York: North Holland Publishing.

Gaudry, M. et al. (2002). ***TRIO, Version 2.0***. Agora Jules Dupuit, Université de Montréal. www.e-ajd.net.

Gaudry, M., Blum, U. and T. Liem (2000). Turning Box-Cox, including Quadratic Forms in Regression. Ch. 14 in Gaudry, M. and S. Lassarre (eds). ***Structural Road Accident Models: The International DRAG Family***, Pergamon, Elsevier Science, Oxford, 335-346.

Gaudry, M. and R. Laferrière (1989). The Box-Cox Transformation: Power Invariance and a New Interpretation. ***Economics Letters*** 30, 27-29.

Gaudry, M. and M. J. Wills (1978). Estimating the Functional Form of Travel Demand Models. ***Transportation Research*** 12, 4, 257-289.

Gaudry, M., and E. Quinet (2003). ***Rail track wear-and-tear costs by traffic class in France***. Publication AJD-66, Agora Jules Dupuit, Université de Montréal, www.e-ajd.net, et W.P. 04-02, Centre d'Enseignement et de Recherche en Analyse Socio-économique, École nationale des Ponts et Chaussées (ENPC-CERAS), www.enpc.fr/ceras/labo/accueil.html, 26 pages, November.

Gaudry, M. and E. Quinet (2009). CATRIN (Cost Allocation of TRansport INfrastructure cost), Deliverable 8, Rail Cost Allocation for Europe – Annex 1Di– ***Track Maintenance Costs in France***. Funded by Sixth Framework Programme. VTI, Stockholm.

Gosset, W. S. ["Student"] (1908). The probable error of a mean. ***Biometrica*** 6, 1, 1-25.

Griffin, R. C., Montgomery, J. M. and M. E. Rister (1987). Selecting Functional Form in Production Function Analysis. ***Western Journal of Agricultural Economics*** 12, 2, 216-227.

Hariton, G. (1984). ***Railway Costing: A Review***. WP 60-64-17, Research Branch, Canadian Transport Commission, 112 pages, October.

Heggie, I. (1991a). ***User Charging and Accountability for Roads: An Agenda for Reform.*** The World Bank, 119 pages, September.

Heggie, I. (1991b). ***Improving Management and Charging Policies for Roads: An Agenda for Reform.*** Report INU 92, The World Bank, 81 pages, December.

Highway Research Board (HRB), (1962). ***The AASHO Road Test, Report 5: Pavement Research***. Special Report 61E.

Hutchinson, B. G. (1991). Some Technical Issues in Implementing Road User Pavement Damage Charges. 18 pages, ***Infrastructure Pricing, Investment Decisions and Financing***, Royal Commission on National Passenger Transportation, Technical Seminar Documentation, Ottawa, 26-27 September.

Jansson, J. O. (2002). An Analysis of the Rail Transport System. In Nash, C., Wardman, M., Button, K. and P. Nijkamp (eds), ***Classics in Transport Analysis: Railways***, 135-181, Edward Elgar Publishing Limited, Cheltenham.

Jara-Díaz, S. (1982). The estimation of transport cost functions: a methodological review. ***Transport Reviews*** 2, 3, pp. 257-278.
Johansson, P. and J.-E. Nilsson (2004). An Economic Analysis of Track Maintenance Costs. ***Transport Policy*** 11, 277–286.
Jung, F. W., Kher, R. and W. A. Phang (1975). ***Ontario Pavement Analysis of Costs: A Performance Prediction System***. Research Report 200, Ministry of Transportation and Communications of Ontario, Downsview, Ontario.
Khaled, M. S. (1978). ***Productivity Analysis and Functional Specification. A Parametric Approach.*** Ph. D. Thesis, Department of Economics, University of British Columbia, April.
Laferrière, R. (1999). ***L'usage du coefficient de Pearson comme mesure du R^2 dans les modèles non linéaires***. Publication AJD-57, Agora Jules Dupuit, Université de Montréal, 2 pages, www.e-ajd.net.
LCPC et SETRA (1994). ***Conception et Dimensionnement des Structures de Chaussée: Guide Technique.*** Laboratoire Central des Ponts et Chaussées et Service d'Études Techniques des Routes et Autoroutes, Ministère de l'Équipement, des Transports et du Tourisme.
Liem. T. C., Dagenais, M. and M. Gaudry (1987). ***The L-1.4 program for BC-GAUHESEQ (Box-Cox Generalized AUtoregressive HEteroskedastic Single EQuation) models***. Publication CRT-510, Centre de recherche sur les transports, Université de Montréal, 41 p., www.e-ajd.net.
Liem, T. C., Gaudry, M., Dagenais, M. and U. Blum (2000). LEVEL: The L-1.5 program for BC-GAUHESEQ (Box-Cox Generalized AUtoregressive HEteroskedastic Single EQuation) regression and multimoment analysis. In Gaudry, M. and S. Lassarre (eds), ***Structural Road Accident Models: The International DRAG Family***, Elsevier Science Publishers, Oxford, Ch. 12, 263-324. www.e-ajd.net.
Link, H. (2009). CATRIN (Cost Allocation of TRansport INfrastructure cost), Deliverable 8, Rail Cost Allocation for Europe – Annex 1C - ***Marginal costs of rail maintenance in Austria***. Funded by Sixth Framework Programme. VTI, Stockholm, February, 23 pages.
Link, H., Stuhlemmer, A. Haraldsson, M., Abrantes, P., Wheat, P., Iwnicki, S., Nash, C., and A. Smith (2008). CATRIN (Cost Allocation of TRansport INfrastructure cost), Deliverable D 1, ***Cost allocation Practices in the European Transport Sector***. Funded by Sixth Framework Programme. VTI, Stockholm, March , 93p.
Lloyd, W. and R. Addis (2000). ***Current Road-Vehicle Interaction Research in the UK***. Paper presented at the International Seminar on Heavy Vehicle Weights and Dimensions, June 18-22, Saskatoon, 12 pages, June.
Marti, M., Neuenschwander, R. and P. Walker (2009). CATRIN (Cost Allocation of TRansport INfrastructure cost) Deliverable 8, Rail Cost Allocation for Europe – Annex 1B–. ***Track maintenance and renewal costs in Switzerland***. Funded by Sixth Framework Programme. VTI, Stockholm.
Martland, C. D. (2001). ***Allocating the Costs of Railroad Infrastructure to Specific Traffic Classes***. Publication AJD-36, Agora Jules Dupuit, Université de Montréal, 20 pages, January 19, 2001. www.e-ajd.net.
Nördling, W. von (1886). Prix de revient des transports par chemin de fer. ***Annales des Ponts et Chaussées*** 11, 10, 292-303.
Olsen, R. J. (1978). ***The method of Box and Cox, a pitfall.*** Institution for Social and Policy Studies, Yale University, Working Paper No 804.
Oum, T. H. and W. G. Waters II (1996). A Survey of Recent Developments in Transportation Cost Function Research. ***Logistics and Transportation Review*** 32, 4, 423-463.
Oum, T. H. and W. G. Waters II (1997). Recent developments in cost function research in transportation. In De Rus, G. and C. Nash (eds.), ***Recent developments in transport economics***, pp. 33–73, Ashgate Publishing, Aldershot, Hampshire.
Poirier, D. J. (1978). The use of the Box-cox transformation in limited dependent variable models. ***Journal of the American Statistical Association*** 73, 284-287.
Quinet, E. (2002). ***Détermination des coûts marginaux sociaux du rail***. Manuscrit, 38 pages, mars.
Rivier, R. et Y. Putallaz (2005). ***Audit sur l'état du réseau ferré national français.*** Laboratoire d'Intermodalité des Transports et de Planification, École Polytechnique fédérale de Lausanne, 30 pages, 7 septembre.
Robert, W. E., Chapman, J. D. and C. D. Martland (1997). Effects of Axle Loads and Train Capacity on Heavy Haul Network Performance. ***Proceedings of the 39th Meeting of the Transportation Research Forum***, Volume 2, 24 pages, October.

Rosett, R. N. and F. D. Nelson (1975). Estimation of the two-limit Probit regression model. ***Econometrica*** 43, 141-146.

Sánchez, P. C. (2000). Vertical Relationships between Railway Operations and the Infrastructure for the European Case. In Nash C. and E. Niskanen (eds), Helsinki Workshop on Infrastructure Charging on Railways, VATT Discussion Papers 245, Government Institute for Economic Research, Helsinki. Pages 71-94.

Schlesselman, J. (1971). Power families: a note on the Box and Cox transformation. ***Journal of the Royal Statistical Society*** Series B 33, 307-31.

Shephard, R. W. (1953). ***Cost and Production Functions.*** Princeton University Press.

Small, K. A. (1990). Highway Infrastructure: Crisis in finance or Crisis in Management? ***ITS Review*** 13, 2, 4-8.

Small, K. A., Winston, C. and C. A. Evans (1989). ***Road Works.*** The Brookings Institution, 128 p., Washington.

Small, K. A. and F. Zhang (1988). ***A Reanalysis of the AASHO Road Test Data: Rigid Pavements***. Department of Economics and Institute of Transportation Studies, University of California, Irvine, 34 pages.

Spitzer, J. J. (1976). The demand for money, the liquidity trap and functional forms. ***International Economic Review*** 17, 220-227.

Spitzer, J. J. (1984). Variance estimates in models with the Box-Cox transformations: Implications for estimation and hypothesis testing. ***Review of Economics and Statistics*** 66, 645-652.

Thomas, J. (2002). ***EU Task Force on Rail Infrastructure charging: summary findings on best practice in marginal cost pricing***. 16 p. IMPRINT–EUROPE Thematic Network (2001-2004), www.imprint-eu.org/public/Papers/IMPRINT3_Thomas.pdf.

Tobin, J. (1958). Estimation of Relationships for Limited Dependent Variables. ***Econometrica*** 25, 24-36.

Tovar, B., Jara-Díaz, S. and L. Trujillo (2003). ***Production and cost functions and their application to the port sector: a literature survey***. World Bank Policy Research Working Paper 3123, 32 pages, August.

UIC, Union Internationale des Chemins de Fer (1989). ***Classification des voies des lignes au point de vue de la maintenance de la voie***. Fiche n° 714 R, 7 pages, Paris, 3ème édition, 1er janvier.

Wheat, P. and S. J. Smith (2008). Assessing the Marginal Infrastructure Maintenance Wear and Tear Costs for Britains's Railway Network. ***Journal of Transport Economics and Policy*** 42, 2, 189-224.

Wheat, P., Smith, A. and C. Nash (2009). CATRIN (Cost Allocation of TRansport INfrastructure cost), ***Deliverable 8 - Rail Cost Allocation for Europe***. Funded by Sixth Framework Programme. VTI, Stockholm.

Wilson, G. W. (1980). ***Economic Analysis of Intercity Freight Transportation***. Indiana University Press, Bloomington, Indiana.

Winston, C. (1985). Conceptual Developments in the Economics of Transportation: An Interpretative Survey***. Journal of Economic Litterature*** 23, pp. 57-94, March.

Zhang, F. (1989). ***A Reanalysis of the AASHO Road Test Data: Flexible Pavements***. Manuscript, Department of Economics and Institute of Transportation Studies, University of California, Irvine, 34 pages.

9. Appendix 1. Interactions among Qualities and Traffics

In the framework of maintenance cost function which is the subject of this study, quality parameters appear as longer term decision parameters, and play a role similar to the role played by capital in cost functions of economic textbooks where the representative function is written:

$$C(T) = f(T,K). \qquad (1)$$

Usual properties of f(T, K) are that this function is increasing in T, *i.e.* ($f'_T > 0$), but U-shaped in K, namely decreasing in K ($f'_K < 0$) for low values of T and increasing in K ($f'_K > 0$) for high values of T. It is easy to deduce that the optimal level of K is such that $f'_K = 0$. Furthermore, it is also often assumed that the marginal short run cost f'T decreases with K, formally stated as $f''_{TK} < 0$ whereby the short run marginal cost decreases with higher capital and equipment levels.

We transpose this property to the case of quality parameters in the maintenance cost function, and assume that:

$$f''_{TQ} < 0. \qquad (2)$$

Let us introduce this assumption in a GBC specification of the maintenance cost function similar to the one presented in the second section, but assuming for simplicity that there is just one traffic and one quality:

$$C^{(\lambda)} = a + b * T^{\alpha} + c * Q^{\beta} + d * T^{\gamma} * Q^{\delta} \qquad (3)$$

What conditions can be derived for the parameters from the assumption that $f''_{TK} < 0$? If $\lambda \neq 1$, no clear condition can be drawn, but if $\lambda=1$, it is easy to see from straightforward calculation that this assumption implies that $\gamma\delta < 0$.

This prompts us to introduce the interaction terms between traffics T and qualities Q through ratios T/Q, the simplest specification ensuring the above result. We therefore use this ratio form for *Maximum allowed speed* and for *Proportion of concrete sleepers* which are parameter directly related to quality, and we use the product for parameters which are inversely correlated to quality: *Age of rails* and *Age of sleepers.*

10. Appendix 2. Effect of quality adjustment on marginal costs

Consider a track on which 2 types of traffic are running and define:

> *C (Q1, Q2, q)* as the ***maintenance*** cost of such a track with two traffics (1 and 2) and the quality of which is measured by the variable *q*;
>
> *F (Q1, Q2, q)* as the ***transport*** cost of these traffics 1 and 2, including monetary, time and quality of track.

Assume that the first derivatives of *C* and *F* with respect to *Q1* and *Q2* and of *C* with respect to q are positive, but that the derivative of *F* with respect to q is negative. Assume further that the Infrastructure Manager aims at achieving a social optimum, a sensible assumption for a public agency, which implies a minimization of the ***total cost*** with respect to the quality level *q*, namely:

Minimize, with respect to q, total cost equal to (C+F).

The first order conditions may then be written as:

$$C'q+F'q=0, \qquad (1)$$

and solved for the optimal value of *q* for given values of *Q1* and *Q2,* yielding a function:

$$q=q\ (Q1,\ Q2). \qquad (2)$$

Then the marginal cost of traffic 1 is:

$$C'\ (Q1) = C'Q1 + (C'q)*(dq/dQ1) \qquad (3)$$

and similarly for traffic 2.

It turns out that the ratio of marginal costs *(dC/dQ1)/(dC/dQ2)* generally differs from the ratio of technical equivalence coefficients *(∂C1/∂Q1)/(∂C/∂Q2)*, which is the ratio of costs to repair damages just *ex-ante*, *i.e.* resetting quality at its *ex-ante* level. This result implies some paradoxical results such as the possibility for a traffic to be very damaging from a technical point of view but have a low marginal cost.

To see it in an extreme case, write simple functions satisfying the above assumptions:

$$C=C_0+C_1*Q1+0,5*C_2*Q1*q^2$$
$$F=F_1*Q1+F_2*Q2-F_3*Q2*q^2$$

which have the built-in property that traffic 1 is damaging but not sensitive to quality, while traffic 2 is not damaging but sensitive to quality, as one expects from freight and passenger traffics. Trivial calculations lead to the following results: $q= (F_3/C_2)*(Q2/Q1)$ and, from (3), to the marginal costs:

$$C'\ (Q1) =C_1-C_2\ (F_3*Q2/C_2*Q1)^2$$
$$C'\ (Q2) =C_2*(Q1/Q2)*(F_3*Q2/C_2*Q1)^2$$

It is clear that, depending on the values of the parameters C_1, C_2 and F_3, *C' (Q1)* and *C' (Q2)* can take any value; and it may happen that *C' (Q1) <C' (Q2)* though traffic 2 causes no damage to the track. The above simple model assumes that the Infrastructure Manager aims at maximising social welfare, but other objectives are possible such as maintaining a given quality level, or making the quality proportionate to the traffic level, or any other option. In those cases, the marginal costs would exhibit different relationships from those derived here.

11. Appendix 3. Results: 5-Form, 4-Quality, 4-Traffic Full models under Rule Z-1

Appendix 3. Beginning (page 1 of 3)

Series I.4.A from Table 8		Linear	Log-Log	Trans-Log	R-GBC	U-GBC
PART I. ELASTICITIES, CONDITIONAL *t*-STATISTICS ($\beta = 0$) and BOX-COX LAMBDA GROUPS						
S: STATE OF LINE SEGMENTS		[X]	[X]	[X]	[X]	[X]
1	Number of switches	.451 (18.12)	.284 λ1 (10.42)	.259 λ1 (8.92)	.283 λ1 (12.63)	.244 λ1 (10.51)
2	Electrified line *Dummy*	.031 (0.49)	.296 (4.91)	.205 (3.09)	.086 (1.57)	.104 (1.98)
3	Automatic switch control *Dummy*	.169 (2.57)	.135 (2.03)	.096 (1.40)	.067 (1.17)	.072 (1.33)
4	Segment length	-.291 (-9.98)	-.298 λ1 (-11.93)	-.290 λ1 (-11.00)	-.270 λ1 (-11.69)	-.255 λ1 (-12.57)
5	Number of tracks = 1 (*vs* 2) *Dummy*	-.055 (-0.79)	-.134 (-1.92)	-.058 (-0.77)	.083 (1.29)	-.003 (-0.06)
6	Number of tracks = 3 (*vs* 2) *Dummy*	-.118 (-0.60)	.263 (1.40)	.080 (0.42)	.055 (0.33)	.116 (0.85)
7	Number of tracks = 4 (*vs* 2) *Dummy*	.259 (2.20)	.573 (5.81)	.370 (3.45)	.289 (3.06)	.279 (2.89)
8	Number of tracks = 5 (*vs* 2) *Dummy*	-1.237 (-3.48)	.155 (0.47)	-.112 (-0.34)	.003 (0.01)	-.157 (-0.40)
9	Number of tracks = 6 (*vs* 2) *Dummy*	.385 (1.44)	.751 (3.50)	.494 (2.20)	.541 (1.87)	.636 (2.85)
10	Number of tracks = 10,18 (*vs* 2) *D.*	.054 (0.10)	1.051 (2.28)	.816 (1.73)	.463 (0.43)	.627 (0.02)
Q: QUALITIES OF SEGMENTS		[X]	[X]	[X]	[X]	[X]
11	Age of rails	.056 (0.65)	.063 λ1 (0.73)	4.598 λ1 (2.01)	.482 λ1 (1.84)	.100 λ1 (1.19)
12	Age of sleepers	.036 (0.41)	.024 λ1 (0.25)	-4,321 λ1 (-1.83)	-.360 λ1 (-1.28)	.052 λ1 (0.66)
13	Maximum speed allowed	-1.017 (-8.24)	-.459 λ1 (-3.88)	-3.652 λ1 (-0.91)	-.050 λ1 (-0.19)	-.482 λ1 (-4.42)
14	Proportion of concrete sleepers	-.126 (-2.46)	.002 λ1 (0.36)	-.050 λ1 (-1.09)	.002 λ1 (0.08)	.017 λ1 (1.53)
Q*Q: QUALITY INTERACTIONS				[ln(X)]*[ln(X)]		
15	(Age of rails)*(Age of rails)			-2.292 (-1.47)		
16	(Age of sleepers)*(Age of sleepers)			.622 (0.36)		
17	(Max. speed)*(Max. speed)			4.689 (0.53)		
18	(% concr.sleepers)*(% concr. sleep.)			-.056 (-1.18)		
19	(Age of rails)*(Age of sleepers)			.765 (0.29)		
20	(Age of rails)*(Max. speed)			-8.39 (-1.22)		
21	(Age of rails)*(% concr. sleepers)			.028 (0.63)		
22	(Age of sleepers)*(Max. speed)			11.561 (1.65)		
23	(Age of sleep.)*(% concr. sleepers)			-.026 (-0.56)		
24	(Max. speed)*(% concr. sleepers)			-.003 (-0.17)		

Continued…

Appendix 3. Continuation (page 2 of 3)

	Series I.4.A from Table 8	Linear	Log-Log	Trans-Log	R-GBC		U-GBC	
PART I. ELASTICITIES, CONDITIONAL *t*-STATISTICS ($\beta = 0$) and BOX-COX LAMBDA GROUPS								
Q*T : QUALITY-TRAFFIC INTER.				[ln(X)]*[ln(X)]	[X]*[X]		[X]*[X][62]	
25	(Age of rails)*(Intercity train t.)			.068 (0.05)	-.056 (-0.20)	λ1	-.026 (-0.43)	λ2
26	(Age of rails)*(Regional train t.)			.122 (0.09)	-.345 (-1.10)	λ1	-.016 (-0.35)	λ2
27	(Age of rails)*(Ile-de-Fr. train t.)			-.013 (-0.10)	-.243 (-1.15)	λ1	.071 (1.95)	λ2
28	(Age of rails)*(Freight train t.)			-2.386 (-2.37)	.053 (0.22)	λ1	.109 (3.13)	λ2
29	(Age of sleepers)*(Intercity train t.)			-.176 (-0.13)	.127 (0.42)	λ1	-.004 (-0.56)	λ3
30	(Age of sleepers)*(Regional train t.)			.622 (0.47)	.470 (1.42)	λ1	.019 (3.08)	λ3
31	(Age of sleepers)*(Ile-de-Fr. train t.)			.005 (0.04)	.440 (2.15)	λ1	.002 (0.55)	λ3
32	(Age of sleepers)*(Freight train t.)			-.040 (-0.04)	-.120 (-0.45)	λ1	.009 (2.32)	λ3
33	(Max. speed)*(Intercity train t.)			.319 (0.38)	-.022 (-0.11)	λ1	.010 (1.70)	λ4
34	(Max. speed)*(Regional train t.)			-.812 (-0.90)	.011 (0.06)	λ1	-.008 (-0.76)	λ4
35	(Max. speed)*(Ile-de-Fr. train t.)			.089 (0.65)	.279 (2.72)	λ1	-.001 (-1.10)	λ4
36	(Max. speed)*(Freight train t.)			2.811 (3.36)	.713 (3.02)	λ1	.025 (3.79)	λ4
37	(% concr. sleep.)*(Intercity train t.)			-.008 (-0.82)	.043 (1.75)	λ1	.001 (2.40)	λ5
38	(% concr. sleep.)*(Regional train t.)			-.002 (-0.21)	.005 (0.21)	λ1	.001 (1.67)	λ5
39	(% concr. sleep.)*(Ile-de-Fr. train t.)			-.017 (-0.76)	.009 (0.34)	λ1	.000 (0.02)	λ5
40	(% concr. sleep.)*(Freight train t.)			.003 (0.39)	-.032 (-1.34)	λ1	-.000 (-1.19)	λ5
T*T : TRAFFIC INTERACTIONS				[ln(X)]*[ln(X)]	[X]*[X]		[X]*[X]	
41	(Intercity train t.)*(Intercity train t.)			.181 (0.77)				
42	(Regional train t.)*(Regional train t.)			.317 (1.86)				
43	(Ile-de-Fr. train t.)*(Ile-de-Fr. train t.)			.208 (2.75)				
44	(Freight train t.)*(Freight train t.)			.291 (1.71)				
45	(Intercity train t.)*(Regional train t.)			-.084 (-0.34)	.070 (1.79)	λ1	-.013 (-0.38)	λ6
46	(Intercity train t.)*(Ile-de-Fr. train t.)			.099 (1.66)	.017 (0.49)	λ1	-.058 (-2.86)	λ6
47	(Intercity train t.)*(Freight train t.)			-.065 (-0.24)	.071 (1.31)	λ1	.019 (0.65)	λ6
48	(Regional train t.)*(Ile-de-Fr. train t.)			-.007 (-0.12)	.005 (0.19)	λ1	.003 (0.26)	λ6
49	(Regional train t.)*(Freight train t.)			.061 (0.22)	-.071 (-1.54)	λ1	-.014 (-0.53)	λ6
50	(Ile-de-Fr. train t.)*(Freight train t.)			-.168 (-2.41)	-.085 (-3.48)	λ1	-.030 (-3.00)	λ6

Continued…

[62] Except for *Maximum allowed speed* (lines 33-36) and *Percent of concrete sleepers* (lines 37-40) which are the denominators of ratio forms, not of products.

Appendix 3. End (p. 3 of 3)

	Series I.4.A from Table 8	Linear	Log-Log	Trans-Log	R-GBC	U-GBC
	PART I. ELASTICITIES, CONDITIONAL *t*-STATISTICS (β = 0) and BOX-COX LAMBDA GROUPS					
	T: TRAFFIC CLASSES	[X]	[X]	[X]	[X]	[X]
51	Intercity train tons	.162 (8.26)	.013 **λ1** (1.53)	-.006 **λ1** (-0.34)	-.048 **λ1** (-0.22)	.111 **λ8** (1.59)
52	Regional train tons	.176 (8.36)	.019 **λ1** (2.79)	.016 **λ1** (1.26)	-.003 **λ1** (-0.01)	.130 **λ7** (2.15)
53	Ile-de-France train tons	.084 (10.06)	.012 **λ1** (4.47)	.011 **λ1** (3.14)	-.307 **λ1** (-2.01)	.059 **λ7** (2.89)
54	Freight train tons	.069 (3.18)	.016 **λ1** (3.14)	.011 **λ1** (1.08)	-.495 **λ1** (-1.69)	-.004 **λ9** (-2.58)
55	REGRESSION CONSTANT	--- (7.28)	--- (27.16)	--- (2.13)	--- (3.27)	--- (10.22)
	PART II. BOX-COX TRANSFORMATIONS, UNCONDITIONAL *t*-STATISTICS [β = 0],[β = 1]					
0	LAMBDA on Cost per km (**λy**)	1.000 FIXED	.000 FIXED	.000 FIXED	.298 [23.55],[-55.37]	.308 [19.90],[-44.75]
1	LAMBDA (Group **λX1**)		.000 FIXED	.000 FIXED	.298 [23.55],[-55.37]	.112 [1.39],[-11.06]
2	LAMBDA (Group **λX2**)					.574 [2.55],[-1.90]
3	LAMBDA (Group **λX3**)					.004 [0.03],[-9.39]
4	LAMBDA (Group **λX4**)					2.129 [3.29],[1.75]
5	LAMBDA (Group **λX5**)					1.594 [2.00],[0.75]
6	LAMBDA (Group **λX6**)					.743 [3.24],[-1.12]
7	LAMBDA (Group **λX7**)					1.114 [4.12],[0.42]
8	LAMBDA (Group **λX8**)					.377 [1.39],[-2.31]
9	LAMBDA (Group **λX9**)					3.458 [1.68],[1.20]
	PART III. GENERAL STATISTICS					
1	LOG-LIKELIHOOD	-13547.6	-13129.3	-13092.5	-12906.4	-12866.4
2	PSEUDO-R2 –(E) –(L) –(E) ADJ. for D.F. –(L) ADJ. for D.F.	.727 .720 .722 .715	.651 .882 .644 .880	.676 .891 .657 .884	.752 .926 .741 .922	.823 .932 .813 .928
3	AVER. PROB. (y = LIMIT OBS.)[63]	.161	.000	.000	.000	.000
4	NUMBER OF OBSERVATIONS	967	967	967	967	967
5	PARAMETERS ESTIMATED BETA Coefficients BOX-COX Transformations	19 0 0	19 0 0	55 55 0	42 41 1	51 41 10

End.

[63] The average value of the first and third right-hand side terms of Equation (8) is calculated.

12. Appendix 4. Dummies associated with variables containing observed zeroes

To understand the role of dummy variables associated to variables that contain zero values but are subjected to a BCT, start with a simple linear regression model with dependent variable ***y*** regressed on a constant ***C***, on a variable ***Q*** that is not strictly positive, *e.g.* snowfall, and on its associated binary variable ***D*** (where $D_t = 1$ if $Q_t \neq 0$ and $D_t = 0$ otherwise):

$$y_t = \beta_0 C_t + \beta Q_t + \beta_d D_t + u_t \,. \tag{1}$$

In (1), the variable $\boldsymbol{D_t}$ can be understood as a variable of interaction capturing the qualitative effect of a *state-of-the world* (there is snow) on $\boldsymbol{y_t}$ beyond the quantitative effect of $\boldsymbol{Q_t}$ (how much snow), or as a test of whether the "first" unit of snowfall has the same impact as the next units. Replacing the state variable by its complement ($D_t = 0$ if $Q_t \neq 0$ and $D_t = 1$ otherwise) simply reverses the sign of β_d and adding a state variable consumes a degree of freedom, modifies the meaning of the model and might improve the adjustment, or not.

But, if the model is instead:

$$y_t = \beta_0 C_t + \beta Q_t^{(\lambda)} + \beta_d D_t + u_t \,, \tag{2}$$

where the BCT is applied only to the positive values of $\boldsymbol{Q_t}$ (to avoid numerical problems with 0 observations when $\lambda \to 0$), the associated binary variable will play an additional role: it will preserve the invariance of the $\boldsymbol{\lambda}$ estimate to changes in the units (scale) of measurement of $\boldsymbol{Q_t}$. The analytic proof of this affirmation is beyond our reach but a numerical proof is accessible if we have OLS regression software: (i) set $\boldsymbol{\lambda}$ in (2) equal to an arbitrary value, such as 2, apply the OLS estimator while conserving $\boldsymbol{Q_t}$ in original units and obtain the vector of coefficients [$\hat{\beta}_0$, $\hat{\beta}$, $\hat{\beta}_d$]; (ii) apply scale factor ***s*** to $\boldsymbol{Q_t}$ and re-estimate with $\tilde{Q}_t = sQ_t$ to obtain [$\tilde{\beta}_0$, $\tilde{\beta}$, $\tilde{\beta}_d$]; (iii) observe that $(\hat{\beta}_0 = \tilde{\beta}_0)$, $(\hat{\beta} = s^{\lambda}\tilde{\beta})$, $(\hat{\beta}_d = \tilde{\beta}_d + s^{(\lambda)}\tilde{\beta})$ and that the error sum of squares is invariant. The model (2) is invariant because the coefficient of the constant $\boldsymbol{C_t}$ does not change, the coefficient of $\boldsymbol{Q_t}$ adjusts in inverse proportion to the factor $\boldsymbol{s^{\lambda}}$ and the scale change compensation (needed by the absence of transformation applied to the zero observations of rescaled $\boldsymbol{Q_t}$) is effected by the coefficient of the associated Boolean dummy variable which has acquired a second role: it still represents a *state-of-the world*, as in (1), but also functions as a scale compensation buffer that makes $\boldsymbol{\lambda}$ invariant to the rescaling of $\boldsymbol{Q_t}$.

We must not confuse this numerical proof of the invariance of $\boldsymbol{\lambda}$ in (2) with the invariance of the $\boldsymbol{\lambda}$ of a BCT applied to a strictly positive ***Q*** variable, demonstrated by Schlesselman (1971) to hold only if the regression has a constant ***C***. In that case, the coefficient of the required constant ***C*** provides the necessary adjustment, after a change of scale ***s*** of the transformed variable ***Q***, as demonstrated in the following simple before and after sequence:

$$\beta_0 C + \beta (Q^{\lambda} - 1)/\lambda) \tag{3-1}$$

$$\beta_0 C + \beta s^{\lambda} (Q/s)^{\lambda} / \lambda - \beta/\lambda \tag{3-2}$$

$$\beta_0 C + \tilde{\beta}\tilde{Q}^{\lambda}/\lambda - \beta/\lambda + \tilde{\beta}/\lambda - \tilde{\beta}/\lambda \tag{3-3}$$

$$\beta_0 C + \tilde{\beta}(\tilde{Q}^{\lambda} - 1)/\lambda - \beta/\lambda + \tilde{\beta}/\lambda \tag{3-4}$$

$$\tilde{\beta}_0 C + \tilde{\beta}(\tilde{Q}^{\lambda} - 1)/\lambda \tag{3-5}$$

where, after the rescaling of ***Q***, its new coefficient $\tilde{\beta} = s^{\lambda}\beta$, the same as in (2), and the new coefficient of the required constant $\tilde{\beta}_0 = (\beta_0 - \beta/\lambda + \tilde{\beta}/\lambda)$, have been *analytically* derived.